

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:46:42 ; Search time 15 Seconds
(without alignments)
351.114 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 179

Sequence: 1 OHNGVLGPYIGKSLTLKP.....ADYTTKATANLYGLNLYRF 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	7.3	77	2	US-08-472-172-12
2	13	7.3	432	2	US-08-472-172-6
3	13	7.3	455	2	US-08-472-172-4
4	13	7.3	459	2	US-08-472-172-2
5	10	5.6	34	1	US-08-475-989-8
6	10	5.6	34	2	US-08-475-989-8
7	10	5.6	34	2	US-08-472-172-20
8	10	5.6	34	3	US-08-256-839-8
9	10	5.6	35	1	US-08-475-989-14
10	10	5.6	35	2	US-08-475-989-14
11	10	5.6	35	2	US-08-472-172-26
12	10	5.6	35	3	US-08-256-839-14
13	10	5.6	38	1	US-08-475-989-13
14	10	5.6	38	2	US-08-475-989-13
15	10	5.6	38	2	US-08-472-172-25
16	10	5.6	38	3	US-08-256-839-13
17	10	5.6	53	1	US-08-475-989-48
18	10	5.6	53	2	US-08-475-989-48
19	10	5.6	53	3	US-08-256-839-48
20	10	5.6	54	1	US-08-475-989-49
21	10	5.6	54	2	US-08-475-989-49
22	10	5.6	54	3	US-08-256-839-49
23	10	5.6	55	1	US-08-475-989-47
24	10	5.6	55	2	US-08-475-989-47
25	10	5.6	55	3	US-08-256-839-47
26	7	3.9	12	4	US-09-347-926-9
27	7	3.9	59	4	US-09-300-672-7

28	7	3.9	61	2	US-08-949-603-11	Sequence 11, Appl
29	7	3.9	61	2	US-08-706-270A-11	Sequence 11, Appl
30	7	3.9	61	2	US-08-949-580-11	Sequence 11, Appl
31	7	3.9	61	2	US-08-950-172A-11	Sequence 11, Appl
32	7	3.9	61	4	US-09-198-119C-11	Sequence 11, Appl
33	7	3.9	68	3	US-08-912-272-28	Sequence 28, Appl
34	7	3.9	68	4	US-09-026-039-28	Sequence 28, Appl
35	7	3.9	69	3	US-08-912-272-18	Sequence 18, Appl
36	7	3.9	69	4	US-09-026-039-18	Sequence 18, Appl
37	7	3.9	466	2	US-08-726-136-28	Sequence 28, Appl
38	7	3.9	466	4	US-09-103-434-28	Sequence 28, Appl
39	7	3.9	466	4	US-09-687-594-28	Sequence 28, Appl
40	6	3.4	15	1	US-08-056-200-60	Sequence 60, Appl
41	6	3.4	15	1	US-08-056-200-61	Sequence 61, Appl
42	6	3.4	15	2	US-08-800-644-60	Sequence 60, Appl
43	6	3.4	15	2	US-08-800-644-61	Sequence 61, Appl
44	6	3.4	15	2	US-08-912-272-7	Sequence 7, Appl
45	6	3.4	18	4	US-09-026-039-7	Sequence 7, Appl
46	6	3.4	20	3	US-08-467-023-40	Sequence 40, Appl
47	6	3.4	20	3	US-08-467-023-41	Sequence 41, Appl
48	6	3.4	21	4	US-09-288-143-181	Sequence 181, Appl
49	6	3.4	22	4	US-08-973-801-15	Sequence 15, Appl
50	6	3.4	26	1	US-08-484-635-91	Sequence 91, Appl
51	6	3.4	26	2	US-08-484-631-91	Sequence 91, Appl
52	6	3.4	26	2	US-08-827-570-91	Sequence 91, Appl
53	6	3.4	30	3	US-08-256-747C-44	Sequence 44, Appl
54	6	3.4	30	3	US-08-467-023-68	Sequence 68, Appl
55	6	3.4	30	3	US-08-467-023-75	Sequence 75, Appl
56	6	3.4	30	4	US-08-834-130A-44	Sequence 44, Appl
57	6	3.4	32	1	US-08-473-989-10	Sequence 10, Appl
58	6	3.4	32	2	US-08-475-985-10	Sequence 10, Appl
59	6	3.4	32	2	US-08-472-172-22	Sequence 22, Appl
60	6	3.4	32	3	US-08-256-839-10	Sequence 10, Appl
61	6	3.4	50	3	US-08-467-023-63	Sequence 63, Appl
62	6	3.4	53	3	US-08-912-272-12	Sequence 12, Appl
63	6	3.4	53	4	US-09-026-039-12	Sequence 12, Appl
64	6	3.4	60	4	US-09-300-672-9	Sequence 9, Appl
65	6	3.4	68	2	US-08-700-152A-2	Sequence 2, Appl
66	6	3.4	68	3	US-08-912-272-5	Sequence 5, Appl
67	6	3.4	68	4	US-09-026-039-5	Sequence 5, Appl
68	6	3.4	69	3	US-08-912-272-17	Sequence 17, Appl
69	6	3.4	69	4	US-09-026-039-17	Sequence 17, Appl
70	6	3.4	90	4	US-09-188-930-134	Sequence 134, App
71	6	3.4	149	4	US-09-227-357-136	Sequence 196, App
72	6	3.4	175	4	US-09-288-143-173	Sequence 173, App
73	6	3.4	245	4	US-09-413-814-6	Sequence 6, Appl
74	6	3.4	253	4	US-09-288-143-84	Sequence 84, Appl
75	6	3.4	262	3	US-08-946-914-14	Sequence 14, Appl
76	6	3.4	262	4	US-09-656-450-14	Sequence 14, Appl
77	6	3.4	264	1	US-08-562-311-4	Sequence 4, Appl
78	6	3.4	276	4	US-08-961-083-134	Sequence 134, App
79	6	3.4	282	4	US-09-288-143-172	Sequence 172, App
80	6	3.4	297	4	US-08-765-907A-3	Sequence 3, Appl
81	6	3.4	299	1	US-08-169-950-3	Sequence 3, Appl
82	6	3.4	333	4	US-09-338-420-4	Sequence 4, Appl
83	6	3.4	359	4	US-09-134-001C-4842	Sequence 4842, Ap
84	6	3.4	374	3	US-08-467-023-2	Sequence 2, Appl
85	6	3.4	386	4	US-09-338-420-2	Sequence 2, Appl
86	6	3.4	393	2	US-08-933-750C-11	Sequence 11, Appl
87	6	3.4	393	4	US-09-234-613-11	Sequence 11, Appl
88	6	3.4	411	4	US-09-188-930-302	Sequence 302, App
89	6	3.4	432	2	US-08-700-152A-4	Sequence 4, Appl
90	6	3.4	446	1	US-07-627-539G-7	Sequence 7, Appl
91	6	3.4	452	3	US-09-059-522-3	Sequence 3, Appl
92	6	3.4	452	4	US-09-382-027-3	Sequence 3, Appl
93	6	3.4	468	1	US-07-627-539G-2	Sequence 2, Appl
94	6	3.4	496	4	US-09-252-149B-26	Sequence 26, Appl
95	6	3.4	496	4	US-08-833-876-2	Sequence 2, Appl
96	6	3.4	496	4	US-09-483-054-2	Sequence 2, Appl
97	6	3.4	503	4	US-09-134-001C-4214	Sequence 4214, Ap
98	6	3.4	507	4	US-09-134-001C-5071	Sequence 5071, Ap
99	6	3.4	512	2	US-08-828-596-2	Sequence 2, Appl
100	6	3.4	512	4	US-09-360-197-6	Sequence 6, Appl

101	6	3.4	524	4	US-09-134-001C-4315	Sequence 4315, Ap	174	6	3.4	908	2	US-08-478-435-94	Sequence 94, Appl
102	6	3.4	533	4	US-09-350-268-3	Sequence 3, Appli	175	6	3.4	908	2	US-08-337-483-94	Sequence 94, Appl
103	6	3.4	533	4	US-09-558-679-3	Sequence 3, Appli	176	6	3.4	908	2	US-08-478-373-94	Sequence 94, Appl
104	6	3.4	544	1	US-08-387-156-10	Sequence 10, Appl	177	6	3.4	908	3	US-08-474-671-94	Sequence 94, Appl
105	6	3.4	544	2	US-08-694-865-10	Sequence 10, Appl	178	6	3.4	908	3	US-08-483-577A-94	Sequence 94, Appl
106	6	3.4	544	2	US-08-878-748-10	Sequence 10, Appl	179	6	3.4	908	3	US-08-448-194-4	Sequence 4, Appl
107	6	3.4	544	3	US-09-124-491-10	Sequence 10, Appl	180	6	3.4	908	3	US-08-613-009A-16	Sequence 16, Appl
108	6	3.4	546	1	US-08-487-890A-99	Sequence 99, Appl	181	6	3.4	908	4	US-08-897-438-94	Sequence 94, Appl
109	6	3.4	546	2	US-08-478-435-99	Sequence 99, Appl	182	6	3.4	908	4	US-08-867-921-4	Sequence 4, Appl
110	6	3.4	546	2	US-08-337-483-99	Sequence 99, Appl	183	6	3.4	908	4	US-08-637-654-94	Sequence 94, Appl
111	6	3.4	546	3	US-08-478-373-99	Sequence 99, Appl	184	6	3.4	908	4	US-08-649-518-94	Sequence 94, Appl
112	6	3.4	546	3	US-08-474-671-99	Sequence 99, Appl	185	6	3.4	908	4	US-08-778-508-22	Sequence 22, Appl
113	6	3.4	546	3	US-08-483-577A-99	Sequence 99, Appl	186	6	3.4	908	4	US-09-059-584-22	Sequence 22, Appl
114	6	3.4	546	4	US-08-897-438-99	Sequence 99, Appl	187	6	3.4	909	2	US-08-363-124A-4	Sequence 4, Appl
115	6	3.4	546	4	US-08-637-654-99	Sequence 99, Appl	188	6	3.4	914	3	US-08-985-916-16	Sequence 16, Appl
116	6	3.4	546	4	US-08-649-518-99	Sequence 99, Appl	189	6	3.4	924	3	US-08-619-812-8	Sequence 8, Appl
117	6	3.4	547	1	US-07-961-522-2	Sequence 2, Appli	190	6	3.4	926	1	US-07-908-253-2	Sequence 2, Appli
118	6	3.4	547	1	US-07-961-522-5	Sequence 5, Appli	191	6	3.4	926	1	US-08-455-970A-2	Sequence 2, Appli
119	6	3.4	547	1	US-08-217-438-2	Sequence 2, Appli	192	6	3.4	926	1	US-08-387-156-6	Sequence 6, Appli
120	6	3.4	547	1	US-08-217-438-6	Sequence 6, Appli	193	6	3.4	926	1	US-08-694-865-6	Sequence 6, Appli
121	6	3.4	547	1	US-08-321-978-2	Sequence 2, Appli	194	6	3.4	926	2	US-08-878-748-6	Sequence 6, Appli
122	6	3.4	547	1	US-08-321-978-5	Sequence 5, Appli	195	6	3.4	926	2	US-08-535-837-2	Sequence 2, Appli
123	6	3.4	547	2	US-08-710-584-2	Sequence 2, Appli	196	6	3.4	926	3	US-09-124-491-6	Sequence 6, Appli
124	6	3.4	547	2	US-08-710-584-5	Sequence 5, Appli	197	6	3.4	926	3	5476657-3	Patent No. 5476657
125	6	3.4	553	4	US-09-422-936-79	Sequence 79, Appl	198	6	3.4	927	3	US-08-895-601-6	Sequence 6, Appli
126	6	3.4	575	1	US-08-403-866-7	Sequence 7, Appli	199	6	3.4	936	1	US-08-455-970A-12	Sequence 12, Appl
127	6	3.4	588	2	US-08-620-605D-2	Sequence 2, Appli	200	6	3.4	936	1	US-08-455-970A-10	Sequence 10, Appl
128	6	3.4	588	2	US-09-005-232A-2	Sequence 2, Appli	201	6	3.4	951	1	US-08-455-970A-14	Sequence 14, Appl
129	6	3.4	600	3	US-08-537-361B-9	Sequence 9, Appli	202	6	3.4	960	4	US-09-422-936-45	Sequence 45, Appl
130	6	3.4	602	3	US-08-990-470A-3	Sequence 3, Appli	203	6	3.4	961	4	US-09-422-936-49	Sequence 49, Appl
131	6	3.4	602	4	US-08-817-707-9	Sequence 9, Appli	204	6	3.4	964	4	US-09-422-936-53	Sequence 53, Appl
132	6	3.4	614	1	US-08-652-207A-2	Sequence 2, Appli	205	6	3.4	977	1	US-08-387-156-8	Sequence 8, Appli
133	6	3.4	621	3	US-09-059-522-1	Sequence 1, Appli	206	6	3.4	977	2	US-08-694-865-8	Sequence 8, Appli
134	6	3.4	621	4	US-09-382-027-1	Sequence 1, Appli	207	6	3.4	977	2	US-08-878-748-8	Sequence 8, Appli
135	6	3.4	622	2	US-08-664-646A-2	Sequence 2, Appli	208	6	3.4	977	3	US-09-124-491-8	Sequence 8, Appli
136	6	3.4	622	2	US-09-066-285-2	Sequence 2, Appli	209	6	3.4	982	1	US-08-127-499A-1	Sequence 1, Appli
137	6	3.4	622	3	US-09-261-006-2	Sequence 2, Appli	210	6	3.4	992	1	US-08-482-847-1	Sequence 1, Appli
138	6	3.4	622	4	US-08-951-088-2	Sequence 2, Appli	211	6	3.4	1018	1	US-08-408-093-6	Sequence 6, Appli
139	6	3.4	622	4	US-09-609-566-2	Sequence 2, Appli	212	6	3.4	1018	1	US-08-408-420A-6	Sequence 6, Appli
140	6	3.4	624	3	US-09-609-570-2	Sequence 9, Appli	213	6	3.4	1018	3	US-08-714-901-6	Sequence 6, Appli
141	6	3.4	628	1	US-08-257-073-9	Sequence 9, Appli	214	6	3.4	1018	3	US-08-040-741-6	Sequence 6, Appli
142	6	3.4	642	2	US-08-245-511-48	Sequence 48, Appl	215	6	3.4	1069	1	US-07-777-715-9	Sequence 9, Appli
143	6	3.4	642	2	US-08-600-993A-48	Sequence 48, Appl	216	6	3.4	1069	1	US-08-170-126-4	Sequence 4, Appli
144	6	3.4	661	1	US-08-514-014-4	Sequence 4, Appli	217	6	3.4	1089	3	US-08-954-418-4	Sequence 4, Appli
145	6	3.4	661	2	US-08-833-823-4	Sequence 4, Appli	218	6	3.4	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
146	6	3.4	662	1	US-07-841-651-4	Sequence 4, Appli	219	6	3.4	1090	4	US-09-346-237-5	Sequence 5, Appli
147	6	3.4	685	5	PCT-US91-09784-4	Sequence 4, Appli	220	6	3.4	1096	4	US-09-346-237-6	Sequence 6, Appli
148	6	3.4	693	1	US-08-056-200-111	Sequence 111, App	221	6	3.4	1098	1	US-07-777-715-7	Sequence 7, Appli
149	6	3.4	693	1	US-08-056-200-112	Sequence 112, App	222	6	3.4	1098	1	US-08-170-126-2	Sequence 2, Appli
150	6	3.4	693	2	US-08-800-644-111	Sequence 111, App	223	6	3.4	1098	3	US-08-954-418-2	Sequence 2, Appli
151	6	3.4	693	2	US-08-800-644-112	Sequence 112, App	224	6	3.4	1334	6	5476657-1	Patent No. 5476657
152	6	3.4	699	2	US-08-694-865-16	Sequence 16, Appl	225	6	3.4	1375	4	US-09-210-361-4	Sequence 4, Appli
153	6	3.4	699	3	US-09-124-491-16	Sequence 16, Appl	226	6	3.4	1375	4	US-09-740-274-4	Sequence 4, Appli
154	6	3.4	771	4	US-09-090-793-8	Sequence 8, Appli	227	6	3.4	1403	1	US-07-908-253-3	Sequence 3, Appli
155	6	3.4	780	1	US-08-375-709-13	Sequence 13, Appl	228	6	3.4	1403	2	US-08-694-865-17	Sequence 17, Appl
156	6	3.4	780	1	US-08-752-929-13	Sequence 13, Appl	229	6	3.4	1403	2	US-08-535-837-3	Sequence 3, Appli
157	6	3.4	790	3	US-08-537-361B-8	Sequence 8, Appli	230	6	3.4	1438	3	US-09-124-491-17	Sequence 17, Appl
158	6	3.4	790	4	US-08-817-707-6	Sequence 6, Appli	231	6	3.4	1438	4	US-09-209-916-1	Sequence 1, Appli
159	6	3.4	790	4	US-08-817-707-8	Sequence 8, Appli	232	6	3.4	1443	2	US-08-670-707A-39	Sequence 39, Appl
160	6	3.4	791	3	US-08-537-361B-4	Sequence 4, Appli	233	6	3.4	1443	2	US-09-037-601-39	Sequence 39, Appl
161	6	3.4	791	4	US-08-817-707-4	Sequence 4, Appli	234	6	3.4	1443	4	US-09-315-179-39	Sequence 39, Appl
162	6	3.4	792	1	US-08-326-670A-2	Sequence 2, Appli	235	6	3.4	1467	4	US-09-523-656-38	Sequence 38, Appl
163	6	3.4	792	3	US-08-537-361B-2	Sequence 2, Appli	236	6	3.4	1471	1	US-08-683-839B-3	Sequence 3, Appli
164	6	3.4	792	3	US-08-537-361B-6	Sequence 6, Appli	237	6	3.4	1475	3	US-09-007-999-2	Sequence 2, Appli
165	6	3.4	792	3	US-08-990-470A-2	Sequence 2, Appli	238	6	3.4	1475	4	US-09-210-361-2	Sequence 2, Appli
166	6	3.4	792	4	US-08-817-707-2	Sequence 2, Appli	239	6	3.4	1475	4	US-09-740-274-2	Sequence 2, Appli
167	6	3.4	844	4	US-09-422-936-47	Sequence 47, Appl	240	6	3.4	1661	2	US-08-882-083-2	Sequence 2, Appli
168	6	3.4	844	4	US-09-422-936-51	Sequence 51, Appl	241	6	3.4	1661	2	US-08-558-107-2	Sequence 2, Appli
169	6	3.4	884	1	US-08-066-167-2	Sequence 2, Appli	242	6	3.4	1661	4	US-09-243-539-2	Sequence 2, Appli
170	6	3.4	886	4	US-09-422-936-77	Sequence 77, Appl	243	6	3.4	2115	4	US-09-324-867-5	Sequence 5, Appli
171	6	3.4	892	4	US-09-422-936-75	Sequence 75, Appl	244	6	3.4	2133	2	US-08-670-707A-37	Sequence 37, Appl
172	6	3.4	899	4	US-09-422-936-71	Sequence 71, Appl	245	6	3.4	2133	4	US-09-037-601-37	Sequence 37, Appl
173	6	3.4	908	1	US-08-487-890A-94	Sequence 94, Appl	246	6	3.4	2133	4	US-09-315-179-37	Sequence 37, Appl

247	6	3.4	2133	4	US-09-523-656-30	Sequence 30, Appl	320	5	2.8	15	6	5496924-4	Patent No. 5496924
248	6	3.4	2304	4	US-09-324-867-4	Sequence 4, Appl1	321	5	2.8	16	6	5496924-13	Patent No. 5496924
249	6	3.4	2319	1	US-08-212-133A-8	Sequence 8, Appl1	322	5	2.8	16	3	US-09-100-414B-2	Sequence 2, Appl1
250	6	3.4	2319	1	US-08-474-503-6	Sequence 6, Appl1	323	5	2.8	16	3	US-09-100-409A-61	Sequence 61, Appl1
251	6	3.4	2319	2	US-08-670-707A-6	Sequence 6, Appl1	324	5	2.8	16	4	US-09-303-323-2	Sequence 2, Appl1
252	6	3.4	2319	4	US-09-037-601-6	Sequence 6, Appl1	325	5	2.8	17	4	US-09-265-653-18	Sequence 18, Appl1
253	6	3.4	2319	4	US-09-315-179-6	Sequence 6, Appl1	326	5	2.8	18	2	US-08-746-283-19	Sequence 19, Appl1
254	6	3.4	2319	4	US-09-523-656-28	Sequence 28, Appl	327	5	2.8	18	2	US-08-746-257A-17	Sequence 17, Appl1
255	6	3.4	2319	5	PCr-US94-13200-6	Sequence 6, Appl1	328	5	2.8	20	1	US-08-190-788A-148	Sequence 148, App
256	6	3.4	2332	1	US-07-864-004B-4	Sequence 4, Appl1	329	5	2.8	20	1	US-08-383-474B-153	Sequence 153, App
257	6	3.4	2332	1	US-08-251-937A-4	Sequence 4, Appl1	330	5	2.8	20	1	US-08-465-391A-148	Sequence 148, App
258	6	3.4	2332	1	US-08-212-133A-2	Sequence 2, Appl1	331	5	2.8	20	2	US-08-482-142-16	Sequence 16, Appl
259	6	3.4	2332	1	US-08-276-594A-2	Sequence 2, Appl1	332	5	2.8	20	2	US-08-482-142-205	Sequence 205, App
260	6	3.4	2332	1	US-08-474-503-2	Sequence 2, Appl1	333	5	2.8	20	2	US-08-464-538B-148	Sequence 148, App
261	6	3.4	2332	2	US-08-670-707A-2	Sequence 2, Appl1	334	5	2.8	20	2	US-08-463-076B-139	Sequence 139, App
262	6	3.4	2332	4	US-09-037-601-2	Sequence 2, Appl1	335	5	2.8	20	2	US-08-348-353-26	Sequence 26, Appl
263	6	3.4	2332	4	US-09-324-867-3	Sequence 3, Appl1	336	5	2.8	20	2	US-08-465-965-26	Sequence 26, Appl
264	6	3.4	2332	4	US-09-315-179-2	Sequence 2, Appl1	337	5	2.8	20	2	US-08-478-572-16	Sequence 16, Appl
265	6	3.4	2332	4	US-09-523-656-2	Sequence 2, Appl1	338	5	2.8	20	2	US-08-478-572-205	Sequence 205, App
266	6	3.4	2332	5	PCr-US93-03275-4	Sequence 4, Appl1	339	5	2.8	20	2	US-08-466-860-29	Sequence 29, Appl
267	6	3.4	2332	5	PCr-US94-13200-2	Sequence 2, Appl1	340	5	2.8	20	3	US-08-465-966-26	Sequence 26, Appl
268	6	3.4	2343	4	US-09-324-867-2	Sequence 2, Appl1	341	5	2.8	20	3	US-08-472-040A-29	Sequence 29, Appl
269	6	3.4	2351	1	US-08-121-202-2	Sequence 2, Appl1	342	5	2.8	20	3	US-08-840-006-4	Sequence 4, Appl1
270	6	3.4	2351	1	US-08-366-851A-2	Sequence 2, Appl1	343	5	2.8	20	4	US-08-276-776-29	Sequence 29, Appl
271	6	3.4	2351	6	5171844-2	Patent No. 5171844	344	5	2.8	20	4	US-08-471-209-29	Sequence 29, Appl
272	6	3.4	2351	6	5422260-1	Patent No. 5422260	345	5	2.8	20	4	US-08-484-296-16	Sequence 16, Appl
273	6	3.4	2384	3	US-08-936-135-4	Sequence 4, Appl1	346	5	2.8	20	4	US-08-484-296-205	Sequence 205, App
274	6	3.4	2388	3	US-08-936-135-2	Sequence 2, Appl1	347	5	2.8	20	4	US-08-055-006-9	Sequence 9, Appl1
275	6	3.4	2710	2	US-08-568-459A-12	Sequence 12, Appl	348	5	2.8	20	5	PCr-US95-04481-7	Sequence 7, Appl1
276	6	3.4	2710	2	US-08-487-826B-12	Sequence 12, Appl	349	5	2.8	21	2	US-08-466-860-30	Sequence 30, Appl
277	6	3.4	2710	2	US-08-210-288-12	Sequence 12, Appl	350	5	2.8	21	3	US-08-472-040A-30	Sequence 30, Appl
278	6	3.4	3060	2	US-08-487-826B-14	Sequence 14, Appl	351	5	2.8	21	3	US-08-851-843A-175	Sequence 175, App
279	6	3.4	3169	2	US-08-477-451-6	Sequence 6, Appl1	352	5	2.8	21	4	US-08-974-549A-294	Sequence 294, App
280	6	3.4	3218	1	US-08-764-100-27	Sequence 27, Appl	353	5	2.8	21	4	US-08-276-776-30	Sequence 30, Appl
281	6	3.4	4551	3	US-09-320-878-1	Sequence 1, Appl1	354	5	2.8	21	4	US-08-471-209-30	Sequence 30, Appl
282	6	3.4	4813	4	US-09-105-537-31	Sequence 31, Appl	355	5	2.8	21	4	US-08-854-050-175	Sequence 175, App
283	6	3.4	11877	4	US-09-105-537-6	Sequence 6, Appl1	356	5	2.8	21	4	US-09-430-323-175	Sequence 175, App
284	5	2.8	7	2	US-08-350-260A-377	Sequence 377, App	357	5	2.8	22	1	US-08-471-780C-4	Sequence 4, Appl1
285	5	2.8	7	4	US-09-139-802-130	Sequence 130, App	358	5	2.8	22	1	US-08-484-635-141	Sequence 141, App
286	5	2.8	9	4	US-08-660-092-125	Sequence 125, App	359	5	2.8	22	2	US-08-467-282B-4	Sequence 4, Appl1
287	5	2.8	9	4	US-09-160-513-125	Sequence 125, App	360	5	2.8	22	2	US-08-484-631-141	Sequence 141, App
288	5	2.8	9	4	US-09-644-600-27	Sequence 27, Appl	361	5	2.8	22	2	US-08-471-282A-4	Sequence 4, Appl1
289	5	2.8	9	4	US-09-644-600-35	Sequence 35, Appl	362	5	2.8	22	2	US-08-466-710C-4	Sequence 4, Appl1
290	5	2.8	9	4	US-09-644-600-36	Sequence 36, Appl	363	5	2.8	22	2	US-08-827-570-141	Sequence 141, App
291	5	2.8	9	4	US-09-644-600-38	Sequence 38, Appl	364	5	2.8	22	3	US-08-468-739C-4	Sequence 4, Appl1
292	5	2.8	10	1	US-08-299-285-21	Sequence 21, Appl	365	5	2.8	22	5	PCr-US92-08094-5	Sequence 5, Appl1
293	5	2.8	10	2	US-08-985-126-21	Sequence 21, Appl	366	5	2.8	23	1	US-08-141-324-25	Sequence 25, Appl
294	5	2.8	10	3	US-08-840-006-3	Sequence 3, Appl1	367	5	2.8	23	1	US-08-541-902-25	Sequence 25, Appl
295	5	2.8	10	4	US-09-358-020-21	Sequence 21, Appl	368	5	2.8	23	1	US-08-139-862-11	Sequence 11, Appl
296	5	2.8	10	5	PCr-US95-11126-21	Sequence 21, Appl	369	5	2.8	23	1	US-08-139-862-12	Sequence 12, Appl
297	5	2.8	11	6	5196523-14	Patent No. 5196523	370	5	2.8	24	1	US-08-321-356-7	Sequence 7, Appl1
298	5	2.8	11	6	5464750-1	Patent No. 5464750	371	5	2.8	24	1	US-08-484-635-70	Sequence 70, Appl
299	5	2.8	13	3	US-08-486-099-174	Sequence 174, App	372	5	2.8	24	2	US-08-482-142-130	Sequence 130, App
300	5	2.8	13	3	US-08-484-223B-174	Sequence 174, App	373	5	2.8	24	2	US-08-484-631-70	Sequence 70, Appl
301	5	2.8	13	3	US-08-919-597-174	Sequence 174, App	374	5	2.8	24	2	US-08-478-572-130	Sequence 130, App
302	5	2.8	13	3	US-08-475-668A-174	Sequence 174, App	375	5	2.8	24	2	US-08-827-570-70	Sequence 70, Appl
303	5	2.8	13	3	US-08-485-551A-174	Sequence 174, App	376	5	2.8	24	3	US-08-545-196B-59	Sequence 59, Appl
304	5	2.8	13	3	US-08-471-913A-174	Sequence 174, App	377	5	2.8	24	4	US-08-484-296-130	Sequence 130, App
305	5	2.8	13	3	US-09-101-146-47	Sequence 47, Appl	378	5	2.8	25	2	US-08-473-475A-6	Sequence 6, Appl1
306	5	2.8	13	4	US-08-485-264A-174	Sequence 174, App	379	5	2.8	25	4	US-08-913-362-26	Sequence 26, Appl
307	5	2.8	13	4	US-09-082-279B-552	Sequence 552, App	380	5	2.8	25	1	US-08-078-683A-15	Sequence 15, Appl
308	5	2.8	13	4	US-08-474-349A-174	Sequence 174, App	381	5	2.8	28	2	US-08-446-692-38	Sequence 38, Appl
309	5	2.8	13	4	US-09-315-304B-532	Sequence 532, App	382	5	2.8	28	2	US-08-482-142-15	Sequence 15, Appl
310	5	2.8	13	6	5496924-54	Patent No. 5496924	383	5	2.8	28	2	US-08-482-142-33	Sequence 33, Appl
311	5	2.8	14	5	PCr-US91-05177-6	Sequence 6, Appl1	384	5	2.8	28	2	US-08-482-142-126	Sequence 126, App
312	5	2.8	14	5	PCr-US93-07261-9	Sequence 9, Appl1	385	5	2.8	28	2	US-08-482-142-127	Sequence 127, App
313	5	2.8	14	6	5496924-48	Patent No. 5496924	386	5	2.8	28	2	US-08-482-142-128	Sequence 128, App
314	5	2.8	15	1	US-08-139-862-8	Sequence 8, Appl1	387	5	2.8	28	2	US-08-482-142-129	Sequence 129, App
315	5	2.8	15	1	US-08-139-862-14	Sequence 14, Appl	388	5	2.8	28	2	US-08-488-351A-38	Sequence 38, Appl
316	5	2.8	15	2	US-08-592-646A-53	Sequence 53, Appl	389	5	2.8	28	2	US-08-348-353-34	Sequence 34, Appl
317	5	2.8	15	4	US-08-913-362-16	Sequence 16, Appl	390	5	2.8	28	2	US-08-465-965-34	Sequence 34, Appl
318	5	2.8	17	4	US-08-913-362-17	Sequence 17, Appl	391	5	2.8	28	2	US-08-478-572-15	Sequence 15, Appl
319	5	2.8	15	4	US-09-165-422-53	Sequence 53, Appl	392	5	2.8	28	2	US-08-478-572-33	Sequence 33, Appl

393	5	2.8	28	2	US-08-478-572-126	Sequence 126, App	466	5	2.8	42	1	US-08-477-383-23	Sequence 23, Appl
394	5	2.8	28	2	US-08-478-572-127	Sequence 127, App	467	5	2.8	42	1	US-08-487-174-23	Sequence 23, Appl
395	5	2.8	28	2	US-08-478-572-128	Sequence 128, App	468	5	2.8	42	1	US-08-480-750-23	Sequence 23, Appl
396	5	2.8	28	2	US-08-478-572-129	Sequence 129, App	469	5	2.8	42	2	US-08-348-353-35	Sequence 35, Appl
397	5	2.8	28	3	US-08-465-966-34	Sequence 34, Appl	470	5	2.8	42	2	US-08-465-966-35	Sequence 35, Appl
398	5	2.8	28	3	US-08-465-966-34	Sequence 34, Appl	471	5	2.8	42	3	US-08-465-966-35	Sequence 35, Appl
399	5	2.8	28	4	US-08-303-323-37	Sequence 37, Appl	472	5	2.8	44	4	US-09-450-072-76	Sequence 76, Appl
400	5	2.8	28	4	US-08-484-296-15	Sequence 15, Appl	473	5	2.8	44	4	US-09-351-348-76	Sequence 76, Appl
401	5	2.8	28	4	US-08-484-296-33	Sequence 33, Appl	474	5	2.8	45	2	US-08-348-353-6	Sequence 6, Appl
402	5	2.8	28	4	US-08-484-296-126	Sequence 126, App	475	5	2.8	45	2	US-08-348-353-22	Sequence 22, Appl
403	5	2.8	28	4	US-08-484-296-127	Sequence 127, App	476	5	2.8	45	2	US-08-465-966-6	Sequence 6, Appl
404	5	2.8	28	4	US-08-484-296-128	Sequence 128, App	477	5	2.8	45	2	US-08-465-966-22	Sequence 22, Appl
405	5	2.8	28	4	US-08-484-296-129	Sequence 129, App	478	5	2.8	45	3	US-08-465-966-6	Sequence 6, Appl
406	5	2.8	28	4	US-08-484-296-129	Sequence 129, App	479	5	2.8	45	3	US-08-465-966-22	Sequence 22, Appl
407	5	2.8	28	2	US-08-482-142-34	Sequence 34, Appl	480	5	2.8	45	4	US-09-227-357-166	Sequence 166, App
408	5	2.8	28	2	US-08-482-142-153	Sequence 153, App	481	5	2.8	45	4	US-08-466-921-30	Sequence 30, Appl
409	5	2.8	28	2	US-08-482-142-154	Sequence 154, App	482	5	2.8	46	1	US-08-446-692-40	Sequence 40, Appl
410	5	2.8	28	2	US-08-478-572-153	Sequence 34, Appl	483	5	2.8	46	2	US-08-488-351A-40	Sequence 40, Appl
411	5	2.8	28	2	US-08-478-572-153	Sequence 153, App	484	5	2.8	47	4	US-09-227-357-251	Sequence 251, App
412	5	2.8	28	2	US-08-478-572-154	Sequence 154, App	485	5	2.8	48	4	US-09-149-476-367	Sequence 367, App
413	5	2.8	28	4	US-08-484-296-34	Sequence 34, Appl	486	5	2.8	49	4	US-09-004-406C-21	Sequence 21, Appl
414	5	2.8	28	4	US-08-484-296-153	Sequence 153, App	487	5	2.8	51	2	US-08-392-625-30	Sequence 30, Appl
415	5	2.8	28	4	US-08-484-296-154	Sequence 154, App	488	5	2.8	51	2	US-08-466-961A-30	Sequence 30, Appl
416	5	2.8	30	2	US-08-482-142-152	Sequence 152, App	489	5	2.8	51	4	US-09-150-460B-16	Sequence 16, Appl
417	5	2.8	30	2	US-08-478-572-152	Sequence 152, App	490	5	2.8	53	4	US-09-586-563C-6	Sequence 6, Appl
418	5	2.8	30	4	US-08-484-296-152	Sequence 152, App	491	5	2.8	53	4	US-09-586-563C-6	Sequence 6, Appl
419	5	2.8	30	4	US-09-425-638A-102	Sequence 102, App	492	5	2.8	54	4	US-09-605-785-557	Sequence 557, App
420	5	2.8	30	4	US-09-543-004-102	Sequence 102, App	493	5	2.8	54	4	US-08-470-179-11	Sequence 11, Appl
421	5	2.8	30	6	5494663-14	Patent No. 5494663	494	5	2.8	59	2	US-08-637-759B-99	Sequence 99, Appl
422	5	2.8	31	2	US-08-482-142-121	Sequence 121, App	495	5	2.8	59	3	US-08-871-355A-99	Sequence 99, Appl
423	5	2.8	31	2	US-08-482-142-122	Sequence 122, App	496	5	2.8	59	3	US-08-468-011A-23	Sequence 23, Appl
424	5	2.8	31	2	US-08-482-142-123	Sequence 123, App	497	5	2.8	59	3	US-08-513-974B-25	Sequence 25, Appl
425	5	2.8	31	2	US-08-482-142-151	Sequence 151, App	498	5	2.8	59	4	US-08-776-971-20	Sequence 20, Appl
426	5	2.8	31	2	US-08-478-572-121	Sequence 121, App	499	5	2.8	59	4	US-09-236-468A-23	Sequence 23, Appl
427	5	2.8	31	2	US-08-478-572-122	Sequence 122, App	500	5	2.8	60	1	US-09-201-945-99	Sequence 99, Appl
428	5	2.8	31	2	US-08-478-572-123	Sequence 123, App	501	5	2.8	61	3	US-08-078-683A-16	Sequence 16, Appl
429	5	2.8	31	4	US-08-478-572-151	Sequence 151, App	502	5	2.8	61	4	US-08-605-150A-16	Sequence 16, Appl
430	5	2.8	31	4	US-08-484-296-121	Sequence 121, App	503	5	2.8	61	4	US-08-981-392-21	Sequence 21, Appl
431	5	2.8	31	4	US-08-484-296-122	Sequence 122, App	504	5	2.8	62	1	US-09-134-001C-5072	Sequence 5072, App
432	5	2.8	31	4	US-08-484-296-151	Sequence 123, App	505	5	2.8	62	1	US-08-202-047-20	Sequence 20, Appl
433	5	2.8	32	1	US-08-324-301-1	Sequence 151, App	506	5	2.8	62	3	US-08-964-690-20	Sequence 20, Appl
434	5	2.8	32	2	US-08-482-142-124	Sequence 124, App	507	5	2.8	62	4	US-09-134-001C-4525	Sequence 4525, App
435	5	2.8	32	2	US-08-482-142-125	Sequence 125, App	508	5	2.8	63	4	US-09-227-357-213	Sequence 213, App
436	5	2.8	32	2	US-08-482-142-163	Sequence 163, App	509	5	2.8	66	4	US-08-971-089-2	Sequence 2, Appl
437	5	2.8	32	2	US-08-482-142-164	Sequence 164, App	510	5	2.8	66	4	US-09-134-001C-5421	Sequence 5421, App
438	5	2.8	32	2	US-08-478-572-124	Sequence 124, App	511	5	2.8	68	3	US-08-513-974B-311	Sequence 311, App
439	5	2.8	32	2	US-08-478-572-125	Sequence 125, App	512	5	2.8	68	4	US-08-776-971-128	Sequence 128, App
440	5	2.8	32	2	US-08-478-572-163	Sequence 163, App	513	5	2.8	69	4	US-09-149-476-617	Sequence 617, App
441	5	2.8	32	4	US-08-478-572-164	Sequence 164, App	514	5	2.8	70	4	US-09-134-001C-4666	Sequence 4666, App
442	5	2.8	32	4	US-08-484-296-124	Sequence 124, App	515	5	2.8	71	3	US-08-297-395-8	Sequence 8, Appl
443	5	2.8	32	4	US-08-484-296-125	Sequence 125, App	516	5	2.8	72	3	US-09-220-528-90	Sequence 90, Appl
444	5	2.8	32	4	US-08-484-296-163	Sequence 163, App	517	5	2.8	72	3	US-09-100-804-19	Sequence 19, Appl
445	5	2.8	32	4	US-08-484-296-164	Sequence 164, App	518	5	2.8	77	2	US-08-348-353-18	Sequence 18, Appl
446	5	2.8	33	4	US-09-447-125B-26	Sequence 26, Appl	519	5	2.8	77	3	US-08-465-965-18	Sequence 18, Appl
447	5	2.8	36	1	US-08-139-862B-6	Sequence 6, Appl	520	5	2.8	77	4	US-08-465-966-18	Sequence 18, Appl
448	5	2.8	36	4	US-09-106-568E-142	Sequence 142, App	521	5	2.8	77	4	US-09-187-789-25	Sequence 25, Appl
449	5	2.8	38	1	US-08-176-500-46	Sequence 46, Appl	522	5	2.8	77	4	US-09-139-600-20	Sequence 20, Appl
450	5	2.8	38	1	US-08-176-500-123	Sequence 123, App	523	5	2.8	78	2	US-08-645-193B-16	Sequence 16, Appl
451	5	2.8	38	1	US-08-471-052A-46	Sequence 46, Appl	524	5	2.8	80	2	US-08-691-814B-30	Sequence 30, Appl
452	5	2.8	38	1	US-08-471-052A-123	Sequence 123, App	525	5	2.8	82	3	US-08-403-852D-27	Sequence 27, Appl
453	5	2.8	38	1	US-08-189-331-46	Sequence 46, Appl	526	5	2.8	82	3	US-08-510-646B-28	Sequence 28, Appl
454	5	2.8	38	1	US-08-189-331-123	Sequence 123, App	527	5	2.8	82	4	US-09-231-818-27	Sequence 27, Appl
455	5	2.8	38	2	US-08-471-939-46	Sequence 46, Appl	528	5	2.8	83	1	US-07-748-783-2	Sequence 2, Appl
456	5	2.8	38	2	US-08-471-939-123	Sequence 123, App	529	5	2.8	83	1	US-08-166-818-2	Sequence 2, Appl
457	5	2.8	38	2	US-08-471-800-46	Sequence 46, Appl	530	5	2.8	84	3	US-09-013-067A-4	Sequence 4, Appl
458	5	2.8	38	2	US-08-471-800-123	Sequence 123, App	531	5	2.8	86	1	US-07-869-933-27	Sequence 27, Appl
459	5	2.8	38	2	US-08-471-068-46	Sequence 46, Appl	532	5	2.8	86	4	US-09-103-663-27	Sequence 27, Appl
460	5	2.8	38	2	US-08-471-068-123	Sequence 123, App	533	5	2.8	86	5	FCT-US95-17083-10	Sequence 10, Appl
461	5	2.8	40	6	5342615-6	Patent No. 5342615	534	5	2.8	88	2	US-08-816-155B-47	Sequence 47, Appl
462	5	2.8	41	2	US-08-640-847C-41	Sequence 41, Appl	535	5	2.8	88	3	US-08-894-173-64	Sequence 64, Appl
463	5	2.8	41	6	US-08-469-260A-539	Sequence 539, App	536	5	2.8	88	4	US-09-079-587-47	Sequence 47, Appl
464	5	2.8	41	6	5196523-6	Patent No. 5196523	537	5	2.8	88	4	US-09-398-193-64	Sequence 64, Appl
465	5	2.8	42	1	US-08-137-800-23	Sequence 23, Appl	538	5	2.8	91	2	US-08-216-276A-27	Sequence 27, Appl
												US-08-637-759B-60	Sequence 60, Appl

539	5	2.8	91	3	US-08-871-355A-60	Sequence 60, Appl	612	5	2.8	126	2	US-08-427-497E-41	Sequence 41, Appl
540	5	2.8	91	4	US-09-025-596-91	Sequence 91, Appl	613	5	2.8	126	4	US-09-298-731-38	Sequence 38, Appl
541	5	2.8	91	4	US-09-201-945-60	Sequence 61, Appl	614	5	2.8	127	4	US-09-173-300-40	Sequence 40, Appl
542	5	2.8	92	4	US-09-344-529-2	Sequence 2, Appl	615	5	2.8	128	6	5179198-1	Patent No. 5179198
543	5	2.8	92	6	5284931-8	Patent No. 5284931	616	5	2.8	128	6	5521296-1	Patent No. 5521296
544	5	2.8	93	1	US-08-603-010A-36	Sequence 36, Appl	617	5	2.8	129	2	US-08-630-822A-102	Sequence 102, App
545	5	2.8	93	1	US-08-680-726A-36	Sequence 36, Appl	618	5	2.8	129	2	US-09-005-069-102	Sequence 102, App
546	5	2.8	93	4	US-09-092-409-36	Sequence 36, Appl	619	5	2.8	129	4	US-09-258-373-8	Sequence 8, Appl
547	5	2.8	93	4	US-08-469-260A-19	Sequence 19, Appl	620	5	2.8	129	4	US-09-171-156A-51	Sequence 51, Appl
548	5	2.8	94	1	US-08-591-498-4	Sequence 4, Appl	621	5	2.8	130	1	US-08-580-680-1	Sequence 1, Appl
549	5	2.8	95	2	US-08-479-078-23	Sequence 23, Appl	622	5	2.8	130	1	US-08-480-156A-1	Sequence 1, Appl
550	5	2.8	95	3	US-08-946-329A-78	Sequence 78, Appl	623	5	2.8	130	1	US-08-354-961-1	Sequence 1, Appl
551	5	2.8	96	4	US-08-858-207A-402	Sequence 402, App	624	5	2.8	130	2	US-08-581-094-1	Sequence 1, Appl
552	5	2.8	97	2	US-08-341-843B-25	Sequence 25, Appl	625	5	2.8	130	2	US-08-580-665-1	Sequence 1, Appl
553	5	2.8	97	2	US-08-427-497E-30	Sequence 30, Appl	626	5	2.8	130	2	US-08-581-142-1	Sequence 1, Appl
554	5	2.8	98	2	US-08-341-843B-13	Sequence 13, Appl	627	5	2.8	130	2	US-08-944-449-2	Sequence 2, Appl
555	5	2.8	98	2	US-08-427-497E-18	Sequence 18, Appl	628	5	2.8	130	2	US-08-944-449-4	Sequence 4, Appl
556	5	2.8	99	4	US-09-134-001C-4683	Sequence 4683, Ap	629	5	2.8	130	3	US-08-516-859A-105	Sequence 105, App
557	5	2.8	101	4	US-09-149-476-754	Sequence 754, App	630	5	2.8	130	4	US-08-581-103-1	Sequence 1, Appl
558	5	2.8	103	2	US-08-537-811-45	Sequence 45, Appl	631	5	2.8	130	4	US-09-586-472-105	Sequence 105, App
559	5	2.8	103	2	US-08-596-319-31	Sequence 31, Appl	632	5	2.8	130	4	US-09-353-362-2	Sequence 2, Appl
560	5	2.8	103	5	PCT-US95-17083-8	Sequence 8, Appl	633	5	2.8	130	4	US-09-353-362-4	Sequence 4, Appl
561	5	2.8	106	1	US-08-276-852-83	Sequence 83, Appl	634	5	2.8	130	4	US-09-528-706-105	Sequence 105, App
562	5	2.8	106	1	US-08-276-852-85	Sequence 85, Appl	635	5	2.8	130	5	PCT-US94-05442A-1	Sequence 1, Appl
563	5	2.8	106	1	US-08-899-575-83	Sequence 83, Appl	636	5	2.8	131	2	US-08-850-910A-39	Sequence 39, Appl
564	5	2.8	106	1	US-08-899-575-85	Sequence 85, Appl	637	5	2.8	131	2	US-08-850-910A-41	Sequence 41, Appl
565	5	2.8	106	1	US-08-899-575-83	Sequence 83, Appl	638	5	2.8	131	2	US-08-850-910A-46	Sequence 46, Appl
566	5	2.8	106	1	US-08-899-575-85	Sequence 85, Appl	639	5	2.8	131	2	US-08-721-498-8	Sequence 8, Appl
567	5	2.8	106	2	US-08-850-910A-30	Sequence 30, Appl	640	5	2.8	131	4	US-09-134-001C-4906	Sequence 4906, Ap
568	5	2.8	106	5	PCT-US95-08743-83	Sequence 83, Appl	641	5	2.8	134	5	PCT-US93-07213-15	Sequence 15, Appl
569	5	2.8	106	5	PCT-US95-08743-85	Sequence 85, Appl	642	5	2.8	135	2	US-08-647-960-5	Sequence 5, Appl
570	5	2.8	107	3	US-08-599-226-9	Sequence 9, Appl	643	5	2.8	137	4	US-09-036-574-4	Sequence 4, Appl
571	5	2.8	107	4	US-09-125-098-9	Sequence 9, Appl	644	5	2.8	137	4	US-09-367-953B-113	Sequence 113, App
572	5	2.8	108	2	US-08-483-695-42	Sequence 42, Appl	645	5	2.8	139	4	US-08-444-818-173	Sequence 173, App
573	5	2.8	108	2	US-08-652-816A-2	Sequence 2, Appl	646	5	2.8	139	4	US-09-367-953B-112	Sequence 112, App
574	5	2.8	108	2	US-08-652-816A-17	Sequence 17, Appl	647	5	2.8	139	4	US-09-367-953B-114	Sequence 114, App
575	5	2.8	108	2	US-08-652-816A-18	Sequence 18, Appl	648	5	2.8	139	4	US-09-367-953B-115	Sequence 115, App
576	5	2.8	108	2	US-08-652-816A-53	Sequence 53, Appl	649	5	2.8	139	4	US-09-367-953B-116	Sequence 116, App
577	5	2.8	108	2	US-07-965-285-42	Sequence 42, Appl	650	5	2.8	139	4	US-09-367-953B-117	Sequence 117, App
578	5	2.8	108	2	US-08-487-231-42	Sequence 42, Appl	651	5	2.8	139	4	US-09-367-953B-118	Sequence 118, App
579	5	2.8	108	2	US-08-997-080-166	Sequence 166, App	652	5	2.8	140	3	US-09-211-631-13	Sequence 13, Appl
580	5	2.8	108	2	US-08-997-382-166	Sequence 166, App	653	5	2.8	140	4	US-09-265-628-13	Sequence 13, Appl
581	5	2.8	108	4	US-09-095-855-166	Sequence 166, App	654	5	2.8	140	4	US-09-001-141-11	Sequence 11, Appl
582	5	2.8	108	4	US-09-095-855-166	Sequence 42, Appl	655	5	2.8	140	4	US-09-532-803-6	Sequence 6, Appl
583	5	2.8	108	4	US-09-129-030-45	Sequence 45, Appl	656	5	2.8	140	4	US-09-367-953B-119	Sequence 119, App
584	5	2.8	108	4	US-09-324-542-166	Sequence 166, App	657	5	2.8	140	4	US-08-858-207A-468	Sequence 468, App
585	5	2.8	108	4	US-09-205-426-166	Sequence 166, App	658	5	2.8	140	4	US-09-653-403-14	Sequence 14, Appl
586	5	2.8	109	4	US-09-187-859-23	Sequence 23, Appl	659	5	2.8	140	4	US-10-013-784-14	Sequence 14, Appl
587	5	2.8	111	1	US-08-466-886-37	Sequence 37, Appl	660	5	2.8	141	2	US-08-187-186A-5	Sequence 5, Appl
588	5	2.8	111	4	US-08-469-617-37	Sequence 37, Appl	661	5	2.8	141	2	US-08-442-497C-9	Sequence 9, Appl
589	5	2.8	112	2	US-08-933-750C-43	Sequence 43, Appl	662	5	2.8	141	4	US-09-333-033-9	Sequence 9, Appl
590	5	2.8	112	3	US-08-928-361B-10	Sequence 10, Appl	663	5	2.8	142	2	US-08-187-186A-2	Sequence 2, Appl
591	5	2.8	112	3	US-08-928-361B-29	Sequence 29, Appl	664	5	2.8	142	2	US-08-442-497C-2	Sequence 2, Appl
592	5	2.8	112	4	US-09-234-613-43	Sequence 43, Appl	665	5	2.8	142	2	US-08-860-174A-7	Sequence 7, Appl
593	5	2.8	112	4	US-09-399-913-38	Sequence 38, Appl	666	5	2.8	142	2	US-09-333-033-2	Sequence 2, Appl
594	5	2.8	114	4	US-08-936-165A-271	Sequence 271, App	667	5	2.8	142	5	PCT-US94-05186-2	Sequence 2, Appl
595	5	2.8	115	4	US-09-256-000-13	Sequence 13, Appl	668	5	2.8	143	4	US-09-258-373-3	Sequence 3, Appl
596	5	2.8	115	4	US-09-308-246C-2	Sequence 2, Appl	669	5	2.8	143	4	US-08-858-207A-441	Sequence 441, App
597	5	2.8	116	4	US-09-308-003-51	Sequence 51, Appl	670	5	2.8	143	6	5169835-18	Patent No. 5169835
598	5	2.8	117	4	US-08-961-083-128	Sequence 128, App	671	5	2.8	145	3	US-09-030-613-9	Sequence 9, Appl
599	5	2.8	118	1	US-08-326-362-2	Sequence 2, Appl	672	5	2.8	145	4	US-09-451-905-9	Sequence 9, Appl
600	5	2.8	118	4	US-08-858-207A-395	Sequence 395, App	673	5	2.8	147	4	US-09-134-001C-3024	Sequence 3024, Ap
601	5	2.8	119	4	US-08-858-207A-397	Sequence 397, App	674	5	2.8	149	4	US-09-586-875-2	Sequence 2, Appl
602	5	2.8	120	4	US-09-228-986-128	Sequence 128, App	675	5	2.8	149	5	PCT-US95-17083-6	Sequence 6, Appl
603	5	2.8	120	4	US-09-615-192A-393	Sequence 393, App	676	5	2.8	150	4	US-09-058-483-3	Sequence 3, Appl
604	5	2.8	121	4	US-08-341-843B-4	Sequence 4, Appl	677	5	2.8	150	4	US-08-936-165A-514	Sequence 514, App
605	5	2.8	121	4	US-09-171-980-2	Sequence 2, Appl	678	5	2.8	154	4	US-09-247-155-110	Sequence 110, App
606	5	2.8	121	4	US-08-353-942-4	Sequence 4, Appl	679	5	2.8	154	4	US-09-134-001C-5287	Sequence 5287, Ap
607	5	2.8	121	5	PCT-US93-03895-4	Sequence 4, Appl	680	5	2.8	155	4	US-09-149-476-466	Sequence 466, App
608	5	2.8	124	4	US-09-425-638A-46	Sequence 46, Appl	681	5	2.8	155	4	US-09-149-476-603	Sequence 603, App
609	5	2.8	124	4	US-09-543-004-46	Sequence 46, Appl	682	5	2.8	155	6	5494663-8	Patent No. 5494663
610	5	2.8	124	4	US-09-199-637A-347	Sequence 347, App	683	5	2.8	156	2	US-08-500-860A-36	Sequence 36, Appl
611	5	2.8	126	2	US-08-341-843B-36	Sequence 36, Appl	684	5	2.8	156	3	US-08-600-982-30	Sequence 30, Appl

685	5	2.8	156	4	US-09-181-183-6	Sequence 6, Appl	758	5	2.8	174	4	US-09-325-256-2	Sequence 2, Appl
686	5	2.8	156	4	US-09-280-040-6	Sequence 6, Appl	759	5	2.8	175	1	US-08-078-683A-34	Sequence 34, Appl
687	5	2.8	156	4	US-09-277-700-6	Sequence 6, Appl	760	5	2.8	175	1	US-08-551-171-8	Sequence 8, Appl
688	5	2.8	156	4	US-09-134-001C-4234	Sequence 4234, Ap	761	5	2.8	175	3	US-08-902-233-8	Sequence 8, Appl
689	5	2.8	156	5	PCT-US94-10261A-30	Sequence 30, Appl	762	5	2.8	175	4	US-08-913-362-4	Sequence 4, Appl
690	5	2.8	157	4	US-09-180-100-15	Sequence 15, Appl	763	5	2.8	175	4	US-08-913-362-30	Sequence 30, Appl
691	5	2.8	157	5	PCT-US93-02475-8	Sequence 8, Appl	764	5	2.8	175	4	US-09-651-941-13	Sequence 13, Appl
692	5	2.8	157	5	PCT-US93-02475-9	Sequence 9, Appl	765	5	2.8	175	4	US-09-325-256-1	Sequence 1, Appl
693	5	2.8	158	4	US-09-134-001C-5467	Sequence 5467, Ap	766	5	2.8	175	4	US-09-955-597-13	Sequence 13, Appl
694	5	2.8	159	1	US-08-193-182-1	Sequence 1, Appl	767	5	2.8	176	1	US-08-415-751-1	Sequence 1, Appl
695	5	2.8	159	1	US-08-139-862-1	Sequence 1, Appl	768	5	2.8	177	4	US-09-058-483-4	Sequence 4, Appl
696	5	2.8	159	1	US-08-139-862-2	Sequence 2, Appl	769	5	2.8	178	3	US-09-081-180-2	Sequence 2, Appl
697	5	2.8	159	2	US-08-606-143-45	Sequence 45, Appl	770	5	2.8	178	3	US-09-040-786-2	Sequence 2, Appl
698	5	2.8	159	2	US-08-599-895-9	Sequence 9, Appl	771	5	2.8	178	4	US-09-247-155-85	Sequence 85, Appl
699	5	2.8	159	3	US-09-211-290-9	Sequence 9, Appl	772	5	2.8	178	4	US-09-280-839-2	Sequence 2, Appl
700	5	2.8	159	3	US-09-322-676-9	Sequence 9, Appl	773	5	2.8	179	2	US-08-933-750C-28	Sequence 28, Appl
701	5	2.8	159	4	US-09-466-036A-9	Sequence 9, Appl	774	5	2.8	179	4	US-09-234-613-28	Sequence 28, Appl
702	5	2.8	159	4	US-09-134-001C-3216	Sequence 3216, Ap	775	5	2.8	180	2	US-08-483-695-41	Sequence 41, Appl
703	5	2.8	159	6	5342615-3	Patent No. 5342615	776	5	2.8	180	2	US-07-965-285-41	Sequence 41, Appl
704	5	2.8	159	6	5494663-6	Patent No. 5494663	777	5	2.8	180	2	US-08-487-231-41	Sequence 41, Appl
705	5	2.8	161	3	US-08-493-071-17	Sequence 17, Appl	778	5	2.8	180	4	US-09-201-912-41	Sequence 41, Appl
706	5	2.8	161	3	US-08-493-071-20	Sequence 20, Appl	779	5	2.8	181	2	US-08-482-142-193	Sequence 193, App
707	5	2.8	162	3	US-09-205-264-3	Sequence 3, Appl	780	5	2.8	181	2	US-08-482-142-195	Sequence 195, App
708	5	2.8	162	4	US-08-536-165A-471	Sequence 471, App	781	5	2.8	181	2	US-08-482-142-197	Sequence 197, App
709	5	2.8	163	2	US-08-531-439B-2	Sequence 2, Appl	782	5	2.8	181	2	US-08-478-572-193	Sequence 193, App
710	5	2.8	163	4	US-09-348-265-2	Sequence 2, Appl	783	5	2.8	181	2	US-08-478-572-195	Sequence 195, App
711	5	2.8	164	4	US-09-370-838-203	Sequence 203, App	784	5	2.8	181	2	US-08-478-572-197	Sequence 197, App
712	5	2.8	165	2	US-08-356-060A-40	Sequence 40, Appl	785	5	2.8	181	4	US-08-484-296-193	Sequence 193, App
713	5	2.8	165	4	US-08-674-509B-40	Sequence 40, Appl	786	5	2.8	181	4	US-08-484-296-195	Sequence 195, App
714	5	2.8	166	3	US-08-513-974B-312	Sequence 312, App	787	5	2.8	181	4	US-08-484-296-197	Sequence 197, App
715	5	2.8	167	2	US-08-690-849-2	Sequence 2, Appl	788	5	2.8	182	4	US-09-129-030-10	Sequence 10, Appl
716	5	2.8	167	3	US-09-004-053-2	Sequence 2, Appl	789	5	2.8	182	4	US-09-129-030-18	Sequence 18, Appl
717	5	2.8	167	3	US-08-513-974B-361	Sequence 361, App	790	5	2.8	183	4	US-09-134-001C-5336	Sequence 5336, Ap
718	5	2.8	167	3	US-08-513-974B-363	Sequence 363, App	791	5	2.8	184	2	US-08-737-825-10	Sequence 10, Appl
719	5	2.8	167	3	US-08-513-974B-367	Sequence 367, App	792	5	2.8	185	3	US-09-122-443-16	Sequence 16, Appl
720	5	2.8	167	4	US-08-776-971-101	Sequence 101, App	793	5	2.8	185	4	US-09-058-483-10	Sequence 10, Appl
721	5	2.8	167	4	US-08-776-971-107	Sequence 107, App	794	5	2.8	187	3	US-08-493-071-16	Sequence 16, Appl
722	5	2.8	167	4	US-08-828-683A-22	Sequence 22, Appl	795	5	2.8	187	3	US-08-493-071-19	Sequence 19, Appl
723	5	2.8	168	1	US-08-468-853-8	Sequence 8, Appl	796	5	2.8	188	1	US-08-233-389C-3	Sequence 3, Appl
724	5	2.8	168	1	US-08-468-855-8	Sequence 8, Appl	797	5	2.8	188	2	US-08-801-863-3	Sequence 3, Appl
725	5	2.8	168	1	US-08-310-357-8	Sequence 8, Appl	798	5	2.8	188	2	US-08-486-596A-3	Sequence 3, Appl
726	5	2.8	168	1	US-08-468-852-8	Sequence 8, Appl	799	5	2.8	188	2	US-08-160-524A-5	Sequence 5, Appl
727	5	2.8	168	2	US-08-468-857-8	Sequence 8, Appl	800	5	2.8	188	2	US-09-004-713-3	Sequence 3, Appl
728	5	2.8	168	3	US-08-669-408B-4	Sequence 4, Appl	801	5	2.8	188	3	US-09-048-889-3	Sequence 3, Appl
729	5	2.8	168	3	US-09-188-579-85	Sequence 85, Appl	802	5	2.8	190	4	US-07-681-701-17	Sequence 17, Appl
730	5	2.8	168	4	US-08-776-971-129	Sequence 129, App	803	5	2.8	191	4	US-09-082-920-5	Sequence 5, Appl
731	5	2.8	168	4	US-09-315-444-85	Sequence 85, Appl	804	5	2.8	192	1	US-08-086-428B-60	Sequence 60, Appl
732	5	2.8	168	4	US-09-721-362-85	Sequence 85, Appl	805	5	2.8	192	1	US-08-086-428B-61	Sequence 61, Appl
733	5	2.8	168	4	US-09-752-165-99	Sequence 99, Appl	806	5	2.8	192	1	US-08-086-428B-62	Sequence 62, Appl
734	5	2.8	169	2	US-08-946-528-8	Sequence 8, Appl	807	5	2.8	192	1	US-08-086-428B-63	Sequence 63, Appl
735	5	2.8	170	1	US-08-551-171-1	Sequence 1, Appl	808	5	2.8	192	1	US-08-086-428B-64	Sequence 64, Appl
736	5	2.8	170	2	US-09-517-347-4	Sequence 4, Appl	809	5	2.8	192	1	US-08-086-428B-65	Sequence 65, Appl
737	5	2.8	170	2	US-08-544-822-14	Sequence 14, Appl	810	5	2.8	192	1	US-08-086-428B-66	Sequence 66, Appl
738	5	2.8	170	3	US-08-902-233-1	Sequence 1, Appl	811	5	2.8	192	1	US-08-086-428B-67	Sequence 67, Appl
739	5	2.8	170	3	US-09-070-964-14	Sequence 14, Appl	812	5	2.8	192	1	US-08-086-428B-68	Sequence 68, Appl
740	5	2.8	170	3	US-09-186-250-4	Sequence 4, Appl	813	5	2.8	192	1	US-08-086-428B-69	Sequence 69, Appl
741	5	2.8	170	4	US-09-517-347-8	Sequence 8, Appl	814	5	2.8	192	1	US-08-086-428B-70	Sequence 70, Appl
742	5	2.8	170	4	US-09-517-347-8	Sequence 8, Appl	815	5	2.8	192	1	US-08-086-428B-71	Sequence 71, Appl
743	5	2.8	170	4	US-09-129-030-40	Sequence 40, Appl	816	5	2.8	192	1	US-08-086-428B-72	Sequence 72, Appl
744	5	2.8	170	4	US-09-518-232A-4	Sequence 4, Appl	817	5	2.8	192	1	US-08-086-428B-73	Sequence 73, Appl
745	5	2.8	170	4	US-09-518-232A-8	Sequence 8, Appl	818	5	2.8	192	1	US-08-086-428B-74	Sequence 74, Appl
746	5	2.8	170	4	US-09-517-358A-4	Sequence 4, Appl	819	5	2.8	192	1	US-08-086-428B-75	Sequence 75, Appl
747	5	2.8	170	4	US-09-517-358A-8	Sequence 8, Appl	820	5	2.8	192	1	US-08-086-428B-76	Sequence 76, Appl
748	5	2.8	171	1	US-08-551-171-2	Sequence 2, Appl	821	5	2.8	192	1	US-08-086-428B-86	Sequence 86, Appl
749	5	2.8	171	2	US-08-934-959-2	Sequence 2, Appl	822	5	2.8	192	1	US-08-086-428B-87	Sequence 87, Appl
750	5	2.8	171	3	US-08-902-233-2	Sequence 2, Appl	823	5	2.8	192	1	US-08-086-428B-88	Sequence 88, Appl
751	5	2.8	171	3	US-08-551-171-7	Sequence 7, Appl	824	5	2.8	192	1	US-08-086-428B-89	Sequence 89, Appl
752	5	2.8	174	1	US-08-460-806-21	Sequence 21, Appl	825	5	2.8	192	1	US-08-086-428B-90	Sequence 90, Appl
753	5	2.8	174	1	US-08-325-630-21	Sequence 21, Appl	826	5	2.8	192	1	US-08-086-428B-91	Sequence 91, Appl
754	5	2.8	174	3	US-08-902-233-7	Sequence 7, Appl	827	5	2.8	192	1	US-08-440-103-39	Sequence 39, Appl
755	5	2.8	174	4	US-08-913-362-2	Sequence 2, Appl	828	5	2.8	192	1	US-08-440-103-42	Sequence 42, Appl
756	5	2.8	174	4	US-08-913-362-6	Sequence 6, Appl	829	5	2.8	192	1	US-08-440-103-43	Sequence 43, Appl
757	5	2.8	174	4	US-08-913-362-8	Sequence 8, Appl	830	5	2.8	192	1	US-08-440-542-39	Sequence 39, Appl

831	5	2.8	192	1	US-08-440-542-42	Sequence 42, Appl	904	5	2.8	192	5	PCT-US95-10398-76	Sequence 76, Appl
832	5	2.8	192	1	US-08-440-542-43	Sequence 43, Appl	905	5	2.8	192	5	PCT-US95-10398-86	Sequence 86, Appl
833	5	2.8	192	1	US-08-231-368-39	Sequence 39, Appl	906	5	2.8	192	5	PCT-US95-10398-87	Sequence 87, Appl
834	5	2.8	192	1	US-08-231-368-42	Sequence 42, Appl	907	5	2.8	192	5	PCT-US95-10398-88	Sequence 88, Appl
835	5	2.8	192	1	US-08-231-368-43	Sequence 43, Appl	908	5	2.8	192	5	PCT-US95-10398-89	Sequence 89, Appl
836	5	2.8	192	1	US-08-440-210-39	Sequence 39, Appl	909	5	2.8	192	5	PCT-US95-10398-90	Sequence 90, Appl
837	5	2.8	192	1	US-08-440-210-42	Sequence 42, Appl	910	5	2.8	192	5	PCT-US95-10398-91	Sequence 91, Appl
838	5	2.8	192	1	US-08-440-210-43	Sequence 43, Appl	911	5	2.8	193	1	US-08-260-202A-19	Sequence 19, Appl
839	5	2.8	192	2	US-08-468-570-60	Sequence 60, Appl	912	5	2.8	193	1	US-08-017-114-19	Sequence 19, Appl
840	5	2.8	192	2	US-08-468-570-61	Sequence 61, Appl	913	5	2.8	193	3	US-09-041-889-5	Sequence 5, Appl
841	5	2.8	192	2	US-08-468-570-62	Sequence 62, Appl	914	5	2.8	193	3	US-08-837-058-5	Sequence 5, Appl
842	5	2.8	192	2	US-08-468-570-63	Sequence 63, Appl	915	5	2.8	193	3	US-08-503-307-19	Sequence 19, Appl
843	5	2.8	192	2	US-08-468-570-64	Sequence 64, Appl	916	5	2.8	193	4	US-09-082-920-4	Sequence 4, Appl
844	5	2.8	192	2	US-08-468-570-65	Sequence 65, Appl	917	5	2.8	193	4	US-09-609-151A-19	Sequence 19, Appl
845	5	2.8	192	2	US-08-468-570-66	Sequence 66, Appl	918	5	2.8	193	5	PCT-US94-02034-19	Sequence 19, Appl
846	5	2.8	192	2	US-08-468-570-67	Sequence 67, Appl	919	5	2.8	195	1	US-08-249-671A-9	Sequence 9, Appl
847	5	2.8	192	2	US-08-468-570-68	Sequence 68, Appl	920	5	2.8	195	4	US-09-206-903A-8	Sequence 8, Appl
848	5	2.8	192	2	US-08-468-570-69	Sequence 69, Appl	921	5	2.8	195	4	US-09-202-122-8	Sequence 8, Appl
849	5	2.8	192	2	US-08-468-570-70	Sequence 70, Appl	922	5	2.8	195	4	US-09-206-935-23	Sequence 23, Appl
850	5	2.8	192	2	US-08-468-570-71	Sequence 71, Appl	923	5	2.8	195	4	US-09-487-792-9	Sequence 9, Appl
851	5	2.8	192	2	US-08-468-570-72	Sequence 72, Appl	924	5	2.8	197	4	US-09-402-668-2	Sequence 2, Appl
852	5	2.8	192	2	US-08-468-570-73	Sequence 73, Appl	925	5	2.8	198	1	US-08-439-725A-14	Sequence 14, Appl
853	5	2.8	192	2	US-08-468-570-74	Sequence 74, Appl	926	5	2.8	198	1	US-08-324-977-22	Sequence 22, Appl
854	5	2.8	192	2	US-08-468-570-75	Sequence 75, Appl	927	5	2.8	198	1	US-08-441-629-12	Sequence 12, Appl
855	5	2.8	192	2	US-08-468-570-76	Sequence 76, Appl	928	5	2.8	198	2	US-08-384-616-22	Sequence 22, Appl
856	5	2.8	192	2	US-08-468-570-86	Sequence 86, Appl	929	5	2.8	198	2	US-08-867-471-14	Sequence 14, Appl
857	5	2.8	192	2	US-08-468-570-87	Sequence 87, Appl	930	5	2.8	198	2	US-08-438-439C-10	Sequence 10, Appl
858	5	2.8	192	2	US-08-468-570-88	Sequence 88, Appl	931	5	2.8	198	2	US-08-438-439C-20	Sequence 20, Appl
859	5	2.8	192	2	US-08-468-570-89	Sequence 89, Appl	932	5	2.8	198	2	US-08-904-686A-22	Sequence 22, Appl
860	5	2.8	192	2	US-08-468-570-90	Sequence 90, Appl	933	5	2.8	198	3	US-08-703-245-15	Sequence 15, Appl
861	5	2.8	192	2	US-08-468-570-91	Sequence 91, Appl	934	5	2.8	198	3	US-08-718-904-15	Sequence 15, Appl
862	5	2.8	192	2	US-08-290-665A-60	Sequence 60, Appl	935	5	2.8	198	3	US-09-023-082A-14	Sequence 14, Appl
863	5	2.8	192	2	US-08-290-665A-61	Sequence 61, Appl	936	5	2.8	198	3	US-08-776-207-12	Sequence 22, Appl
864	5	2.8	192	2	US-08-290-665A-62	Sequence 62, Appl	937	5	2.8	198	4	US-09-315-850-22	Sequence 5, Appl
865	5	2.8	192	2	US-08-290-665A-63	Sequence 63, Appl	938	5	2.8	198	4	US-09-068-740A-5	Sequence 5, Appl
866	5	2.8	192	2	US-08-290-665A-64	Sequence 64, Appl	939	5	2.8	198	5	PCT-US95-09172-12	Sequence 12, Appl
867	5	2.8	192	2	US-08-290-665A-65	Sequence 65, Appl	940	5	2.8	198	5	PCT-US95-09172-12	Sequence 12, Appl
868	5	2.8	192	2	US-08-290-665A-66	Sequence 66, Appl	941	5	2.8	199	6	5510472-9	Patent No. 5510472
869	5	2.8	192	2	US-08-290-665A-67	Sequence 67, Appl	942	5	2.8	200	4	US-09-298-731-32	Sequence 32, Appl
870	5	2.8	192	2	US-08-290-665A-68	Sequence 68, Appl	943	5	2.8	201	2	US-08-531-525-13	Sequence 13, Appl
871	5	2.8	192	2	US-08-290-665A-69	Sequence 69, Appl	944	5	2.8	201	2	US-08-718-270A-13	Sequence 13, Appl
872	5	2.8	192	2	US-08-290-665A-70	Sequence 70, Appl	945	5	2.8	202	1	US-08-048-700-2	Sequence 2, Appl
873	5	2.8	192	2	US-08-290-665A-71	Sequence 71, Appl	946	5	2.8	202	2	US-08-853-261A-3	Sequence 3, Appl
874	5	2.8	192	2	US-08-290-665A-72	Sequence 72, Appl	947	5	2.8	202	3	US-08-839-711-3	Sequence 3, Appl
875	5	2.8	192	2	US-08-290-665A-73	Sequence 73, Appl	948	5	2.8	202	4	US-09-227-224-3	Sequence 3, Appl
876	5	2.8	192	2	US-08-290-665A-74	Sequence 74, Appl	949	5	2.8	203	4	US-08-855-825-12	Sequence 12, Appl
877	5	2.8	192	2	US-08-290-665A-75	Sequence 75, Appl	950	5	2.8	203	4	US-08-855-825-14	Sequence 14, Appl
878	5	2.8	192	2	US-08-290-665A-76	Sequence 76, Appl	951	5	2.8	204	4	US-09-230-637-25	Sequence 25, Appl
879	5	2.8	192	2	US-08-290-665A-86	Sequence 86, Appl	952	5	2.8	204	4	US-08-508-296B-58	Sequence 58, Appl
880	5	2.8	192	2	US-08-290-665A-87	Sequence 87, Appl	953	5	2.8	204	4	US-09-230-371A-26	Sequence 26, Appl
881	5	2.8	192	2	US-08-290-665A-88	Sequence 88, Appl	954	5	2.8	205	2	US-08-912-227-4	Sequence 4, Appl
882	5	2.8	192	2	US-08-290-665A-89	Sequence 89, Appl	955	5	2.8	205	4	US-08-883-086-8	Sequence 8, Appl
883	5	2.8	192	2	US-08-290-665A-90	Sequence 90, Appl	956	5	2.8	205	4	US-09-589-287B-4	Sequence 4, Appl
884	5	2.8	192	2	US-08-290-665A-91	Sequence 91, Appl	957	5	2.8	205	4	US-09-513-584-4	Sequence 4, Appl
885	5	2.8	192	4	US-09-046-604-39	Sequence 39, Appl	958	5	2.8	206	3	US-09-041-889-6	Sequence 6, Appl
886	5	2.8	192	4	US-09-046-604-42	Sequence 42, Appl	959	5	2.8	206	3	US-08-837-058-6	Sequence 6, Appl
887	5	2.8	192	4	US-09-046-604-43	Sequence 43, Appl	960	5	2.8	206	3	US-08-513-974B-27	Sequence 27, Appl
888	5	2.8	192	5	PCT-US95-10398-60	Sequence 60, Appl	961	5	2.8	206	4	US-08-776-971-22	Sequence 22, Appl
889	5	2.8	192	5	PCT-US95-10398-61	Sequence 61, Appl	962	5	2.8	207	3	US-08-551-171-5	Sequence 5, Appl
890	5	2.8	192	5	PCT-US95-10398-62	Sequence 62, Appl	963	5	2.8	207	3	US-08-902-233-5	Sequence 5, Appl
891	5	2.8	192	5	PCT-US95-10398-63	Sequence 63, Appl	964	5	2.8	207	4	US-09-134-001C-3806	Sequence 3806, Ap
892	5	2.8	192	5	PCT-US95-10398-64	Sequence 64, Appl	965	5	2.8	207	4	US-09-417-721-10	Sequence 10, Appl
893	5	2.8	192	5	PCT-US95-10398-65	Sequence 65, Appl	966	5	2.8	208	1	US-08-551-171-6	Sequence 6, Appl
894	5	2.8	192	5	PCT-US95-10398-66	Sequence 66, Appl	967	5	2.8	208	1	US-08-464-590A-16	Sequence 16, Appl
895	5	2.8	192	5	PCT-US95-10398-67	Sequence 67, Appl	968	5	2.8	208	1	US-08-462-169B-14	Sequence 14, Appl
896	5	2.8	192	5	PCT-US95-10398-68	Sequence 68, Appl	969	5	2.8	208	2	US-08-207-412B-11	Sequence 11, Appl
897	5	2.8	192	5	PCT-US95-10398-69	Sequence 69, Appl	970	5	2.8	208	2	US-08-951-822-27	Sequence 27, Appl
898	5	2.8	192	5	PCT-US95-10398-70	Sequence 70, Appl	971	5	2.8	208	3	US-09-103-079-14	Sequence 14, Appl
899	5	2.8	192	5	PCT-US95-10398-71	Sequence 71, Appl	972	5	2.8	208	3	US-08-902-233-6	Sequence 6, Appl
900	5	2.8	192	5	PCT-US95-10398-72	Sequence 72, Appl	973	5	2.8	208	3	US-09-093-585-16	Sequence 16, Appl
901	5	2.8	192	5	PCT-US95-10398-73	Sequence 73, Appl	974	5	2.8	208	4	US-09-368-951-27	Sequence 27, Appl
902	5	2.8	192	5	PCT-US95-10398-74	Sequence 74, Appl	975	5	2.8	208	4	US-09-710-092-8	Sequence 8, Appl
903	5	2.8	192	5	PCT-US95-10398-75	Sequence 75, Appl	976	5	2.8	209	4	US-09-109-100-8	Sequence 8, Appl

977 5 2.8 209 4 US-09-109-100-9. Sequence 9, Appl
978 5 2.8 209 4 US-09-109-100-11. Sequence 11, Appl
979 5 2.8 209 4 US-09-109-100-12. Sequence 12, Appl
980 5 2.8 209 4 US-09-109-100-13. Sequence 13, Appl
981 5 2.8 209 4 US-09-109-100-14. Sequence 14, Appl
982 5 2.8 209 4 US-09-109-100-15. Sequence 15, Appl
983 5 2.8 209 4 US-09-109-100-16. Sequence 16, Appl
984 5 2.8 209 4 US-09-109-100-17. Sequence 17, Appl
985 5 2.8 209 4 US-09-109-100-18. Sequence 18, Appl
986 5 2.8 209 4 US-08-506-296B-61. Sequence 61, Appl
987 5 2.8 210 4 US-09-897-537A-2. Sequence 2, Appl
988 5 2.8 212 1 US-08-792-019B-9. Sequence 9, Appl
989 5 2.8 212 3 US-08-988-819-9. Sequence 9, Appl
990 5 2.8 212 4 US-09-016-534-9. Sequence 9, Appl
991 5 2.8 212 4 US-08-612-973-4. Sequence 4, Appl
992 5 2.8 212 4 US-08-097-869-7. Sequence 7, Appl
993 5 2.8 212 4 US-08-795-473B-6. Sequence 6, Appl
994 5 2.8 212 4 US-08-927-597-4. Sequence 4, Appl
995 5 2.8 212 4 US-09-230-637-45. Sequence 45, Appl
996 5 2.8 212 4 US-09-155-941-1. Sequence 1, Appl
997 5 2.8 212 4 US-09-109-100-10. Sequence 10, Appl
998 5 2.8 212 4 US-08-702-525-63. Sequence 63, Appl
999 5 2.8 212 4 US-09-230-371A-27. Sequence 27, Appl
1000 5 2.8 212 4 US-09-439-856-6. Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-472-172-12
; Sequence 12, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 12:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 12:
; Sequence 12, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 12:
; Sequence 12, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 12:
; Query Match: 7.3%; Score 13; DB 2; Length 432;
; Best Local Similarity 100.0%; Pred. No. 2.8e-05;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-12
; Query Match: 7.3%; Score 13; DB 2; Length 77;
; Best Local Similarity 100.0%; Pred. No. 5.5e-06;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-12
; QY 121 GATYKFTPNLSVD 133
; Db 15 GATYKFTPNLSVD 27
; RESULT 2
US-08-472-172-6
; Sequence 6, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-6
; QY 121 GATYKFTPNLSVD 133
; Db 15 GATYKFTPNLSVD 27
; Query Match: 7.3%; Score 13; DB 2; Length 432;
; Best Local Similarity 100.0%; Pred. No. 2.8e-05;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-6
; QY 121 GATYKFTPNLSVD 133
; Db 15 GATYKFTPNLSVD 27

Db 390 GATYKFTPNLSVD 402

RESULT 3

```
US-08-472-172-4
; Sequence 4, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerly, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-4
```

Query Match 7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133

|||||

Db 393 GATYKFTPNLSVD 405

RESULT 4

```
US-08-472-172-2
; Sequence 2, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
```

```
; APPLICANT: McVerly, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-2
```

Query Match 7.3%; Score 13; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133

|||||

Db 397 GATYKFTPNLSVD 409

RESULT 5

```
US-08-475-989-8
; Sequence 8, Application US/08475989
; Patent No. 5679352
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-989-8

Query Match 5.6%; Score 10; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
|||||
Db 2 ELSGFHQLTD 11

RESULT 6

US-08-475-985-8
Sequence 8, Application US/08475985
Patent No. 5972349
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R6
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,985
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-985-8

Query Match 5.6%; Score 10; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
|||||
Db 2 ELSGFHQLTD 11

RESULT 7

US-08-472-172-20
Sequence 20, Application US/08472172
Patent No. 5985288
GENERAL INFORMATION:
APPLICANT: Munson, Jr., Robert S
APPLICANT: Grass, Susan
APPLICANT: Chong, Pele Y
APPLICANT: Fahim, Raafat
APPLICANT: Sia, Charles D Y
APPLICANT: McVerry, Patrick
APPLICANT: Klein, Michel
TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
TITLE OF INVENTION: of Haemophilus Influenzae Type B
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R6
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,172
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/849,411
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-471
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163
TELEX: 063-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-472-172-20

Query Match 5.6%; Score 10; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
Db 2 ELSGFHQLTD 11

RESULT 8

US-08-256-839-8
Sequence 8, Application US/08256839
Patent No. 6018019

GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/256.839
FILING DATE:
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-839-8

Query Match 5.6%; Score 10; DB 3; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
Db 2 ELSGFHQLTD 11

RESULT 9

US-08-475-989-14

Sequence 14, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/475.989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424

APPLICATION NUMBER: US 08/256.839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:v9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-989-14

Query Match 5.6%; Score 10; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
Db 24 ANLYGLNLNY 33

RESULT 10

US-08-475-985-14
Sequence 14, Application US/08475985
Patent No. 5972349

GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine

```

; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,985
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA93/00041
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; APPLICATION NUMBER: GB 9202219.3
; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MTS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-475-985-14

Query Match 5.6%; Score 10; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
Db 24 ANLYGLNLNY 33

RESULT 11
US-08-472-172-26
; Sequence 26, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; OF HAEMOPHILUS INFLUENZAE TYPE B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario

```

```

; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-472-172-26

Query Match 5.6%; Score 10; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
Db 24 ANLYGLNLNY 33

RESULT 12
US-08-256-839-14
; Sequence 14, Application US/08256839
; Patent No. 6018019
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: Sia, Charles
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,839
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
; TELECOMMUNICATION INFORMATION:

```


TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-839-14

Query Match 5.6%; Score 10; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
DB 24 ANLYGLNLNY 33
|||||

RESULT 13

US-08-475-989-13
Sequence 13, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993

CLASSIFICATION: 424
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-475-989-13

Query Match 5.6%; Score 10; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
DB 27 ANLYGLNLNY 36
|||||

RESULT 14

US-08-475-985-13
Sequence 13, Application US/08475985
Patent No. 5972349
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,985
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-475-985-13

Query Match 5.6%; Score 10; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
DB 27 ANLYGLNLNY 36
|||||

RESULT 15
US-08-472-172-25
; Sequence 25, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: Of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-25

Query Match 5.6%; Score 10; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 168 ANLYGLNLNY 177
Db 27 ANLYGLNLNY 36

RESULT 16
US-08-256-839-13
; Sequence 13, Application US/08256839
; Patent No. 6018019
; GENERAL INFORMATION:
; APPLICANT: Chong, Pele
; APPLICANT: Kandil, Ali
; APPLICANT: Sia, Charles
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,839
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jlb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-256-839-13

Query Match 5.6%; Score 10; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 168 ANLYGLNLNY 177
Db 27 ANLYGLNLNY 36

RESULT 17
US-08-475-989-48
; Sequence 48, Application US/08475989
; Patent No. 5679352
; GENERAL INFORMATION:
; APPLICANT: Chong, Pele
; APPLICANT: Kandil, Ali
; APPLICANT: Sia, Charles
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,989
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-989-48

Query Match 5.6%; Score 10; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 ANYGLNLY 177
Db 42 ANYGLNLY 51

RESULT 18
US-08-475-985-48
Sequence 48, Application US/08475985
Patent No. 5972349
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,985
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.

REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-985-48

Query Match 5.6%; Score 10; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 ANYGLNLY 177
Db 42 ANYGLNLY 51

RESULT 19
US-08-256-839-48
Sequence 48, Application US/08256839
Patent No. 6018019
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,839
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-839-48

Query Match 5.6%; Score 10; DB 3; Length 53;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 ANYGLNLY 177
Db 42 ANYGLNLY 51

```

RESULT 20
US-08-475-989-49
: Sequence 49, Application US/08475989
: Patent No. 5679352
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: KANDIL, Ali
: APPLICANT: SIA, Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: Synthetic Haemophilus Influenzae
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/475,989
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/256,839
: FILING DATE: 03-FEB-1993
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/CA93/00041
: FILING DATE: 03-FEB-1992
: CLASSIFICATION: 424
: NAME: STEWART, MICHAEL I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 54 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-475-989-49

```

```

Query Match 5.6%; Score 10; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANYGLNLY 177
Db 43 ANYGLNLY 52

```

```

RESULT 21
US-08-475-985-49
: Sequence 49, Application US/08475985
: Patent No. 5972349
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: KANDIL, Ali
: APPLICANT: SIA, Charles

```

```

: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: Synthetic Haemophilus Influenzae
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/475,985
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/256,839
: FILING DATE: 03-FEB-1993
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/CA93/00041
: FILING DATE: 03-FEB-1993
: CLASSIFICATION: 424
: NAME: STEWART, MICHAEL I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 54 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-475-985-49

```

```

Query Match 5.6%; Score 10; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANYGLNLY 177
Db 43 ANYGLNLY 52

```

```

RESULT 22
US-08-256-839-49
: Sequence 49, Application US/08256839
: Patent No. 6018019
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: KANDIL, Ali
: APPLICANT: SIA, Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: Synthetic Haemophilus Influenzae
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario

```

/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/256,839
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEWART, MICHAEL I.
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 54 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-256-839-49

Query Match 5.6%; Score 10; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLY 177
Db 43 ANLYGLNLY 52

RESULT 23
US-08-475-989-47
/ Sequence 47, Application US/08475989
/ Patent No. 5679352
/ GENERAL INFORMATION:
/ APPLICANT: CHONG, Pele
/ APPLICANT: KANDIL, Ali
/ APPLICANT: SIA, Charles
/ APPLICANT: KLEIN, Michel
/ TITLE OF INVENTION: Synthetic Haemophilus Influenzae
/ NUMBER OF SEQUENCES: 56
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: Suite 701, 330 University Avenue
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/475,989
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/256,839
/ FILING DATE: 03-FEB-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/CA93/00041
/ FILING DATE: 03-FEB-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9202219.3

/ FILING DATE: 03-FEB-1992
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEWART, MICHAEL I.
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ INFORMATION FOR SEQ ID NO: 47:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 55 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-475-989-47

Query Match 5.6%; Score 10; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLY 177
Db 44 ANLYGLNLY 53

RESULT 24
US-08-475-985-47
/ Sequence 47, Application US/08475985
/ Patent No. 5972349
/ GENERAL INFORMATION:
/ APPLICANT: CHONG, Pele
/ APPLICANT: KANDIL, Ali
/ APPLICANT: SIA, Charles
/ APPLICANT: KLEIN, Michel
/ TITLE OF INVENTION: Synthetic Haemophilus Influenzae
/ NUMBER OF SEQUENCES: 56
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: Suite 701, 330 University Avenue
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/475,985
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/256,839
/ FILING DATE: 03-FEB-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/CA93/00041
/ FILING DATE: 03-FEB-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9202219.3
/ FILING DATE: 03-FEB-1992
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEWART, MICHAEL I.
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 55 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-475-985-47

Query Match 5.6%; Score 10; DB 2; Length 55;

Best Local Similarity 100.0%; Pred. No. 0.0047;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 ANLYGLNLNY 177

Db 44 ANLYGLNLNY 53

RESULT 25

US-08-256-839-47

; Sequence 47, Application US/08256839

; Patent No. 6018019

; GENERAL INFORMATION:

; APPLICANT: CHONG, Pete

; APPLICANT: KANDIL, Ali

; APPLICANT: SIA, Charles

; APPLICANT: KLEIN, Michel

; TITLE OF INVENTION: Synthetic Haemophilus Influenzae

; TITLE OF INVENTION: Conjugate Vaccine

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: Suite 701, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/256.839

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, MICHAEL I.

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-373 MTS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 55 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-256-839-47

Query Match

Best Local Similarity 5.6%; Score 10; DB 3; Length 55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 ANLYGLNLNY 177

Db 44 ANLYGLNLNY 53

RESULT 26

US-09-347-926-9

; Sequence 9, Application US/09347926

; Patent No. 6440386

; GENERAL INFORMATION:

; APPLICANT: LEUNG, SHUI-ON

; TITLE OF INVENTION: STABILIZED RADIOPHOSPHATE-LABELED PROTEINS

; FILE REFERENCE: 018733/0936

; CURRENT APPLICATION NUMBER: US/09/347.926

; CURRENT FILING DATE: 1999-07-06

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-347-926-9

Query Match

Best Local Similarity 3.9%; Score 7; DB 4; Length 12;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 QDGSQYE 63

Db 2 QDGSQYE 8

RESULT 27

US-09-300-672-7

; Sequence 7, Application US/09300672

; Patent No. 6248937

; GENERAL INFORMATION:

; APPLICANT: Finkelstein, Ruth R.

; APPLICANT: Lynch, Tim

; APPLICANT: Goodman, Howard M.

; APPLICANT: Wang, Ming-Li

; TITLE OF INVENTION: A TRANSCRIPTION FACTOR REGULATING SEED DEVELOPMENT,

; FILE REFERENCE: 480.89(HV)

; CURRENT APPLICATION NUMBER: US/09/300.672

; CURRENT FILING DATE: 1999-04-27

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 59

; TYPE: PRT

; ORGANISM: AP2 domain protein

US-09-300-672-7

Query Match

Best Local Similarity 3.9%; Score 7; DB 4; Length 59;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 LAYDKAA 101

Db 38 LAYDKAA 44

RESULT 28

US-08-949-603-11

; Sequence 11, Application US/08949603

; Patent No. 5891859

; GENERAL INFORMATION:

; APPLICANT: Michael F. Thomas and

; APPLICANT: Eric J. Stockinger

; TITLE OF INVENTION: DNA AND ENCODED PROTEIN

; TITLE OF INVENTION: WHICH REGULATES COLD AND

; TITLE OF INVENTION: DEHYDRATION REGULATED GENES

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ian C. McLeod

; STREET: 2190 Commons Parkway

; CITY: Okemos

; STATE: Michigan

; COUNTRY: USA

; ZIP: 48864

; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS 5.00
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,603
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/706,270
FILING DATE: September 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-384
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5891859e
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 61
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide
HYPOTHETICAL: NO
ANTI-SENSE: No
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Tobacco
STRAIN:
INDIVIDUAL ISOLATE: N/A
DEVELOPMENTAL STAGE: N/A
HAPLOTYPE: N/A
TISSUE TYPE: N/A
CELL TYPE: N/A
CELL LINE: N/A
ORGANELLE: N/A
IMMEDIATE SOURCE: N/A
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: sequencing
OTHER INFORMATION: Figure 2D
PUBLICATION INFORMATION:
US-08-949-603-11

Query Match 3.9%; Score 7; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
DB 40 LAYDKAA 46

RESULT 29
US-08-706-270A-11
Sequence 11, Application US/08706270A
Patent No. 5892009
GENERAL INFORMATION:
APPLICANT: Michael F. Thomashow and
APPLICANT: Eric J. Stockinger
TITLE OF INVENTION: DNA AND ENCODED PROTEIN
WHICH REGULATES COLD AND
DEHYDRATION REGULATED GENES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway

CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS 5.00
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,270A
FILING DATE: September 4, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-310
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5892009e
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 61
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide
HYPOTHETICAL: NO
ANTI-SENSE: No
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Tobacco
STRAIN:
INDIVIDUAL ISOLATE: N/A
DEVELOPMENTAL STAGE: N/A
HAPLOTYPE: N/A
TISSUE TYPE: N/A
CELL TYPE: N/A
CELL LINE: N/A
ORGANELLE: N/A
IMMEDIATE SOURCE: N/A
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: sequencing
OTHER INFORMATION: Figure 2D
PUBLICATION INFORMATION:
US-08-706-270A-11

Query Match 3.9%; Score 7; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
DB 40 LAYDKAA 46

RESULT 30
US-08-949-580-11
Sequence 11, Application US/08949580
Patent No. 5929305
GENERAL INFORMATION:
APPLICANT: Michael F. Thomashow and
APPLICANT: Eric J. Stockinger
TITLE OF INVENTION: DNA AND ENCODED PROTEIN
WHICH REGULATES COLD AND
DEHYDRATION REGULATED GENES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway

;; TITLE OF INVENTION: DEHYDRATION REGULATED GENES
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ian C. McLeod
;; STREET: 2190 Commons Parkway
;; CITY: Okemos
;; STATE: Michigan
;; COUNTRY: USA
;; ZIP: 48864
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS 5.00
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/949,580
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/706,270
;; FILING DATE: September 4, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ian C. McLeod
;; REGISTRATION NUMBER: 20,931
;; REFERENCE/DOCKET NUMBER: MSU 4.1-383
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (517) 347-4100
;; TELEFAX: (517) 347-4103
;; TELEX: NO. 5929305e
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 61
;; TYPE: Amino Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; MOLECULE TYPE: Polypeptide
;; HYPOTHETICAL: No
;; ANTI-SENSE: No
;; FRAGMENT TYPE:
;; ORIGINAL SOURCE:
;; ORGANISM: Tobacco
;; STRAIN:
;; INDIVIDUAL ISOLATE: N/A
;; DEVELOPMENTAL STAGE: N/A
;; HAPLOTYPE: N/A
;; TISSUE TYPE: N/A
;; CELL TYPE: N/A
;; CELL LINE: N/A
;; ORGANELLE: N/A
;; IMMEDIATE SOURCE: N/A
;; POSITION IN GENOME: N/A
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD: sequencing
;; OTHER INFORMATION: Figure 2D
;; PUBLICATION INFORMATION:
US-08-949-580-11

Query Match 3.9%; Score 7; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 95 LAYDKAA 101
|||||
Db 40 LAYDKAA 46

RESULT 31
US-08-950-172A-11
; Sequence 11, Application US/08950172A
; Patent No. 5965705

;; GENERAL INFORMATION:
;; APPLICANT: Michael F. Thomashow and
;; APPLICANT: Eric J. Stockinger
;; TITLE OF INVENTION: DNA AND ENCODED PROTEIN
;; TITLE OF INVENTION: WHICH REGULATES COLD AND
;; TITLE OF INVENTION: DEHYDRATION REGULATED GENES
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ian C. McLeod
;; STREET: 2190 Commons Parkway
;; CITY: Okemos
;; STATE: Michigan
;; COUNTRY: USA
;; ZIP: 48864
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS 5.00
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/950,172A
;; FILING DATE: 10/14/1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/949,580
;; FILING DATE: October 14, 1997
;; APPLICATION NUMBER: 08/706,270
;; FILING DATE: September 4, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ian C. McLeod
;; REGISTRATION NUMBER: 20,931
;; REFERENCE/DOCKET NUMBER: MSU 4.1-383
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (517) 347-4100
;; TELEFAX: (517) 347-4103
;; TELEX: NO. 5965705e
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 61
;; TYPE: Amino Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; MOLECULE TYPE: Polypeptide
;; HYPOTHETICAL: No
;; ANTI-SENSE: No
;; FRAGMENT TYPE:
;; ORIGINAL SOURCE:
;; ORGANISM: Tobacco
;; STRAIN:
;; INDIVIDUAL ISOLATE: N/A
;; DEVELOPMENTAL STAGE: N/A
;; HAPLOTYPE: N/A
;; TISSUE TYPE: N/A
;; CELL TYPE: N/A
;; CELL LINE: N/A
;; ORGANELLE: N/A
;; IMMEDIATE SOURCE: N/A
;; POSITION IN GENOME: N/A
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD: sequencing
;; OTHER INFORMATION: Figure 2D
;; PUBLICATION INFORMATION:
US-08-950-172A-11

Query Match 3.9%; Score 7; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Qy 95 LAYDKAA 101
|||||

Db 40 LAYDKAA 46

RESULT 32

US-09-198-119C-11
; Sequence 11, Application US/09198119C
; Patent No. 6417428
; GENERAL INFORMATION:

APPLICANT: Thomashow, Michael
APPLICANT: Stockinger, Eric
APPLICANT: Jaglo-Octosen, Kirsten
APPLICANT: Gilmour, Sarah
APPLICANT: Zarka, Daniel
APPLICANT: Jiang, Cai-Zhong

FILE REFERENCE: 19117.713 Seq List
TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance

CURRENT APPLICATION NUMBER: US/09/198,119C

CURRENT FILING DATE: 1998-11-23

PRIOR APPLICATION NUMBER: US 08/706,270

PRIOR FILING DATE: 1998-09-04

PRIOR APPLICATION NUMBER: US 09/018,233

PRIOR FILING DATE: 1998-02-03

PRIOR APPLICATION NUMBER: US 09/017,816

PRIOR FILING DATE: 1998-02-03

PRIOR APPLICATION NUMBER: US 09/018,235

PRIOR FILING DATE: 1998-02-03

PRIOR APPLICATION NUMBER: US 09/017,575

PRIOR FILING DATE: 1998-02-03

PRIOR APPLICATION NUMBER: US 09/018,227

PRIOR FILING DATE: 1998-02-03

PRIOR APPLICATION NUMBER: US 09/018,234

PRIOR FILING DATE: 1998-02-03

NUMBER OF SEQ ID NOS: 95

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 11

LENGTH: 61

TYPE: PRT

ORGANISM: Nicotiana tabacum

US-09-198-119C-11

Query Match 3.9%; Score 7; DB 4; Length 61;

Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101

|||||

Db 40 LAYDKAA 46

RESULT 33

US-08-912-272-28

; Sequence 28, Application US/08912272

; Patent No. 6093874

; GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/912,272

FILING DATE: 15-AUG-1997

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067220US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..68
OTHER INFORMATION: /note= "EREBP-like Arabidopsis RAP2.4
OTHER INFORMATION: AP2 domain"
FEATURE:
NAME/KEY: Region
LOCATION: 35..50
OTHER INFORMATION: /note= "putative RAP2.4 amphipathic
OTHER INFORMATION: alpha-helix"
US-08-912-272-28

Query Match 3.9%; Score 7; DB 3; Length 68;

Best Local Similarity 100.0%; Pred. No. 6.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 LAYDKAA 101

|||||

Db 41 LAYDKAA 47

RESULT 34

US-09-026-039-28

; Sequence 28, Application US/09026039

; Patent No. 6329567

; GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/026,039

FILING DATE: 19-FEB-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,272

FILING DATE: 15-AUG-1997

PRIOR APPLICATION DATA: US 08/879,827

APPLICATION NUMBER: US/08/879,827

FILING DATE: 20-JUN-1997

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..68
; OTHER INFORMATION: /note= "putative RAP2.4 amphipathic
; OTHER INFORMATION: alpha-helix"
; US-09-026-039-28
;
; Query Match 3.9%; Score 7; DB 4; Length 68;
; Best Local Similarity 100.0%; Pred. No. 6.9;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 95 LAYDKAA 101
; DB 41 LAYDKAA 47
;
; RESULT 35
; US-08-912-272-18
; Sequence 18, Application US/08912272
; Patent No. 6093874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,272
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION NUMBER: US 08/879,827
; APPLICATION DATA: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067220US
;

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..69
; OTHER INFORMATION: /note= "AP2 domain within tobacco
; OTHER INFORMATION: EREBP-2"
; US-08-912-272-18
;
; Query Match 3.9%; Score 7; DB 3; Length 69;
; Best Local Similarity 100.0%; Pred. No. 7;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 95 LAYDKAA 101
; DB 42 LAYDKAA 48
;
; RESULT 36
; US-09-026-039-18
; Sequence 18, Application US/09026039
; Patent No. 6325567
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA: US 08/879,827
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 18:
;

```

SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..69
OTHER INFORMATION: /note= "AP2 domain within tobacco
ERBP-2"
FEATURE:
NAME/KEY: Region
LOCATION: 35..51
OTHER INFORMATION: /note= "putative ERBP-2 amphipathic
alpha-helix"
US-09-026-039-18

Query Match 3.9%; Score 7; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
DB 42 LAYDKAA 48

RESULT 37
US-08-726-136-28
; Sequence 28, Application US/08726136
; Patent No. 5811286
; GENERAL INFORMATION:
; APPLICANT: ROBERT D. FALLON
; APPLICANT: MARK S. PAYNE
; APPLICANT: MARK J. NELSON
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,136
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/004914
; FILING DATE: OCTOBER 6, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA A.
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9677
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: AMIDASE
; US-09-026-039-18

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: AMIDASE
; US-08-726-136-28

Query Match 3.9%; Score 7; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LRLAGLAY 97
DB 443 LRLAGLAY 449

RESULT 38
US-09-103-434-28
; Sequence 28, Application US/09103434
; Patent No. 6133421
; GENERAL INFORMATION:
; APPLICANT: ROBERT D. FALLON
; APPLICANT: MARK S. PAYNE
; APPLICANT: MARK J. NELSON
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,136
; FILING DATE:
; APPLICATION NUMBER: 60/004914
; FILING DATE: OCTOBER 6, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA A.
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9677
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: AMIDASE
; US-09-103-434-28

Query Match 3.9%; Score 7; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 Lraglay 97
| | | | |
Db 443 Lraglay 449

RESULT 39

US-09-687-594-28
; Sequence 28, Application US/09687594
; Patent No. 6251650
; GENERAL INFORMATION:
; APPLICANT: ROBERT D. FALLON
; APPLICANT: MARK S. PAYNE
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS ENCODING
; TITLE OF INVENTION: STEREOSPECIFIC NITRILE HYDRATASE AND AMIDASE ENZYMES AND
; TITLE OF INVENTION: RECOMBINANT ORGANISMS EXPRESSING THOSE ENZYMES USEFUL FOR
; TITLE OF INVENTION: THE PRODUCTION OF CHIRAL AMIDES AND ACIDS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION NUMBER: US/09/687,594
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/726,136
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA A.
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9677
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: AMIDASE
; US-09-687-594-28

Query Match 3.9%; Score 7; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 Lraglay 97
| | | | |
Db 443 Lraglay 449

RESULT 40

US-08-056-200-60
; Sequence 60, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul

; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-056-200-60

Query Match 3.4%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 NLSVDV 134
| | | | |
Db 7 NLSVDV 12

Search completed: May 12, 2003, 09:49:11
Job time : 30 secs

XX
PT

PT vaccine against infections, esp. in infants and for diagnosis
 PS Disclosure; Fig 5; 33pp: English.
 CC Plasmid PRSM793 contains only the 3' portion of the pl gene. The
 CC plasmid is derived from PRSM188 which contains the full-length pl
 CC sequence. This cro-lacZ-omp1 fusion protein is produced by PRSM793
 CC and is recognised by rabbit and guinea pig pl-specific antisera in
 CC immunoblot analyses.
 CC See also AARI2446-R12455.
 CC
 XX Sequence 77 AA;
 SQ

Query Match 26.8%; Score 256; DB 12; Length 77;
 Best Local Similarity 64.9%; Pred. No. 2.5c-20;
 Matches 50; Conservative 10; Mismatches 13; Indels 4; Gaps 1;
 QY 107 SASIPDTRDMYSIGATYKFTPNLSVDVGFALHGRKKKHFFETQNIKG----LLLVEADY 162
 Db 1 SASIPDTRDMYSIGATYKFTPNLSVDVGFALHGRKKKHFFETQNIKG----LLLVEADY 162
 QY 163 TTKATANLYGLNLYRF 179
 Db 61 TSQAHANLYGLNLYSF 77

RESULT 2
 AAY75757
 ID AAY75757 standard; Protein; 459 AA.
 XX
 AC AAY75757;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria gonorrhoeae ORF 989 protein sequence SEQ ID NO:2586.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 OS
 XX
 PN W0957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 XX WPI: 2000-062150/05.
 DR N-PSDB; AA254519.
 DR
 XX
 XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 XX Claim 2; Page 1404; 1453pp; English.
 PS
 XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 459 AA;
 Query Match 15.8%; Score 151; DB 21; Length 459;
 Best Local Similarity 22.3%; Pred. No. 9.3e-08;
 Matches 43; Conservative 36; Mismatches 74; Indels 40; Gaps 5;
 QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSKYKTEWSRFK-----ELRGKYQDGSYGVEAF 65
 Db 281 KASVKIVTPESLSVHGMVKVSDKADLFGDVTWTRHSRKNKAELEFEKKNANGCKSDRT 340
 QY 66 TKKEEYKDNRSFAIGTYYSLNDALTLRAGLAYDK--AASKTHLSASIPDTRDMYSIGAT 123
 Db 341 TITPNWRNTYKVGGLGGYQISEPLQLRVGIAFDKPPVRNADYRNLSPLDGNRIWFSAGMK 400
 QY 124 YKFTPNLSVDVGFALHGRKKKHFFETQNIKGLLVEADYTTKATAN----- 169
 Db 401 YHICKNHVVDAAAYTHI-----HINDT-----SVRTAKSGNDVDKSGASCARFKN 445
 QY 170 ---LYGLNLYRF 179
 Db 446 HADIIGLYTYKF 458

RESULT 3
 AAB01259
 ID AAB01259 standard; Protein; 442 AA.
 XX
 AC AAB01259;
 XX
 DT 27-OCT-2000 (first entry)
 XX
 DE Neisseria meningitidis BASB044 protein #2.
 XX
 KW Meningitis; microbial disease; upper respiratory tract infection;
 KW bacteraemia; invasive bacterial disease; BASB044 protein.
 XX
 OS Neisseria meningitidis.
 XX
 PN W0200034482-A2.
 XX
 PD 15-JUN-2000.
 XX
 PF 07-DEC-1999; 99WO-IB02014.
 XX
 PR 08-DEC-1998; 98GB-0026979.
 PR 08-DEC-1998; 98GB-0026980.
 PR 17-DEC-1998; 98GB-0028015.
 PR 03-JAN-1999; 99GB-0000090.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 XX Ruelle J, Veriant VGCL;
 PI
 XX WPI: 2000-423427/36.
 DR N-PSDB; AAA49619.
 DR
 XX Novel BASB041, 43, 44 and 48 polypeptides of Neisseria meningitidis
 PT useful for diagnostic, prophylactic and therapeutic purposes against
 PT microbial diseases comprise a specific amino acid sequence
 XX

PS Claim 52; Page 166-167; 171pp; English.

XX The present sequence is the BASB044 protein from the *Neisseria*

CC meningitidis strain H44/76. The protein and its gene can be used in

CC the prevention and treatment of microbial diseases such as bacteraemia,

CC meningitis and upper respiratory tract infections. They are particularly

CC useful for treating bacterial diseases. They can also be used for

CC diagnosing these diseases. The sequence shows strong homology to

CC the BASB044 protein from *N. meningitidis* strain ATCC 13090.

XX

SQ Sequence 442 AA;

Query Match 15.1%; Score 144.5; DB 21; Length 442;

Best Local Similarity 23.4%; Pred. No. 4.5e-07;

Matches 45; Conservative 38; Mismatches 70; Indels 39; Gaps 7;

OY 12 KGSLLKLPAYWELSGFHOLTDQWAIHYSKYKTEWSRF-----KE---LRGKYQDGS 60

DB 265 KARVKIVTPESLSVHGMVKVSKADLFGDVTWTRHSRFDKAEVLPEKEKTVVYKKG----- 319

OY -61 GYEAFTRKEEYKDNSRFAIGTGYSLNDALTLAGLAYDKAASKT--HLSASIPDTRDMY 118

DB 320 -SDRTTITPNWRNTYKVGFGSGVQISEPLQLRAGIAFDKSPVRNADYRMNSLPDGNRIWF 378

OY 119 SIGATYKFTPNLSVDVGFALHGRKKKHFFVETQ-----NIKGLLLVADYTTTKAT 167

DB 379 SAGMKYHIGKNHVDAAVYTHI-----HINDTSYRTAKASGNDVDVSKG-----ASSARFKNH 429

OY 168 ANLYGLNLNRYF 179

DB 430 ADIIGLQYTYKF 441

RESULT 4

AA75758

ID AA75758 standard; Protein; 466 AA.

XX

AC AA75758;

XX

DT 21-MAR-2000 (first entry)

XX

DE *Neisseria meningitidis* ORF 989 protein sequence SEQ ID NO:2988.

XX

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;

KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;

XX antibacterial; gene therapy.

OS *Neisseria meningitidis*.

XX

PN WO9957280-A2.

XX

PD 11-NOV-1999.

XX

PF 30-APR-1999; 99WO-US09346.

XX

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098994.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

XX

PA (CHIR) CHIRON CORP.

PA (GENQ-) INST GENOMIC RES.

XX

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;

PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

PI Tettelin H, Venter JC;

XX

XX WPI; 2000-062150/05.

DR N-PSDB; AA254520.

XX Novel *Neisseria* polypeptides predicted to be useful antigens for

PT vaccines and diagnostics

XX

PS Claim 2; Page 1404-1405; 1453pp; English.

XX

CC AA253015 to AA254536, AA254577 to AA254615, and AA254253 to AA254591

CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides

CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254673 represent

CC PCR primers used in the exemplification of the present invention. The

CC polypeptides, the polynucleotides, antibodies and compositions of

CC the invention can be used as vaccines, as diagnostic reagents, and as

CC immunogenic compositions. The polypeptides can be used in the

CC manufacture of medicaments for treating or preventing infection due to

CC *Neisseria* bacteria (e.g. meningitis and septicaemia), to detect the

CC presence of *Neisseria* bacteria, or to raise antibodies. They may also

CC be used to screen for agonists or antagonists, which may themselves

CC have use as antibacterial agents. The polynucleotides of the invention

CC may also be used in gene therapy protocols.

XX

SQ Sequence 466 AA;

Query Match 15.1%; Score 144.5; DB 21; Length 466;

Best Local Similarity 23.4%; Pred. No. 4.9e-07;

Matches 45; Conservative 38; Mismatches 70; Indels 39; Gaps 7;

OY 12 KGSLLKLPAYWELSGFHOLTDQWAIHYSKYKTEWSRF-----KE---LRGKYQDGS 60

DB 289 KARVKIVTPESLSVHGMVKVSKADLFGDVTWTRHSRFDKAEVLPEKEKTVVYKKG----- 343

OY 61 GYEAFTRKEEYKDNSRFAIGTGYSLNDALTLAGLAYDKAASKT--HLSASIPDTRDMY 118

DB 344 -SDRTTITPNWRNTYKVGFGSGVQISEPLQLRAGIAFDKSPVRNADYRMNSLPDGNRIWF 402

OY 119 SIGATYKFTPNLSVDVGFALHGRKKKHFFVETQ-----NIKGLLLVADYTTTKAT 167

DB 403 SAGMKYHIGKNHVDAAVYTHI-----HINDTSYRTAKASGNDVDVSKG-----ASSARFKNH 453

OY 168 ANLYGLNLNRYF 179

DB 454 ADIIGLQYTYKF 465

RESULT 5

AAB01258

ID AAB01258 standard; Protein; 466 AA.

XX

AC AAB01258;

XX

DT 27-OCT-2000 (first entry)

XX

DE *Neisseria meningitidis* BASB044 protein #1.

XX

KW Meningitis; microbial disease; upper respiratory tract infection;

KW bacteraemia; invasive bacterial disease; BASB044 protein.

OS *Neisseria meningitidis*.

XX

PH Key Location/Qualifiers

FT Peptide 1..24

FT Protein /label= signal_peptide

FT 25..466

FT /label= mature_BASB044

XX

PN WO200034482-A2.

XX

PD 15-JUN-2000.

XX

PF 07-DEC-1999; 99WO-IB02014.

XX

PR 08-DEC-1998; 98GB-0026979.

PR 08-DEC-1998; 98GB-0026980.

PR 17-DEC-1998; 98GB-0028015.

```

PR 05-JAN-1999; 99GB-0000090.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J, Variant VGCL;
XX
XX WPI; 2000-423427/36.
XX N-PSDB; AAA49618.
XX
XX Novel BASB041, 43, 44 and 48 polypeptides of Neisseria meningitidis
XX useful for diagnostic, prophylactic and therapeutic purposes against
XX microbial diseases comprise a specific amino acid sequence
XX
XX Claim 52; Page 163-165; 171pp; English.
XX
XX The present sequence is the BASB044 protein from the Neisseria
XX meningitidis strain ATCC 13090. The protein and its gene can be used in
XX the prevention and treatment of microbial disease such as bacteraemia,
XX meningitis and upper respiratory tract infections. They are particularly
XX useful for treating bacterial diseases. They can also be used for
XX diagnosing these diseases. The protein sequence shows significant
XX homology to the long chain fatty acid transport protein FadL of E. coli.
XX
XX Sequence 466 AA;
XX
XX Query Match 15.0%; Score 143.5; DB 21; Length 466;
XX Best Local Similarity 23.4%; Pred. No. 6.3e-07;
XX Matches 45; Conservative 38; Mismatches 70; Indels 39; Gaps 7;
XX
XX 12 KGSLLTLPAYWELSGFHOLTOWAIHYSYKYTWSRF-----KE---LRGKYODGS 60
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 289 KARVKIYTPESLVHGYKVSADKADLFQDVTWTRHSRFDKAEVFEKEKTVKVGK----- 343
XX
XX 61 GYEAFTRKEEYKDNSRFAICTTYSNLDAITLRAGLAYDKAAKST--HLSASIPDTRMWY 118
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 344 -SDRTTTPNWRNTYKVGCGSVQISEPLQLRAGIAFDKSPVRNADYRMNSLPDGNRIWF 402
XX
XX 119 SIGATYKFTPNLSVDVGFALHGRGKKKHVFETO-----NIKGLLLVEADYTTKAT 167
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 403 SAGMKYHIGKNHVVDAAAYTHI-----HINDTYSRTAKASGNDVDSKG-----ASSAREKNH 453
XX
XX 168 ANLYGLNLNRYF 179
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 454 ADIIGLOQTYTKF 465
XX
XX RESULT 6
XX AAY75759
XX ID AAY75759 standard; Protein; 464 AA.
XX AC AAY75759;
XX
XX 21-MAR-2000 (first entry)
XX
XX Neisseria meningitidis ORF 989 protein sequence SEQ ID NO:2990.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX antibacterial; gene therapy.
XX
XX Neisseria meningitidis.
XX
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
XX 31-JUL-1998; 98US-0094869.
XX 02-SEP-1998; 98US-0098994.
XX 03-SEP-1998; 98US-0099062.
XX 03-OCT-1998; 98US-0103749.
XX
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Piazza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX WPI; 2000-062150/05.
XX N-PSDB; AA254521.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics
XX
XX Claim 2; Page 1406; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AA254253 to AA255941
XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of
XX the invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the
XX manufacture of medicaments for treating or preventing infection due to
XX Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX presence of Neisseria bacteria, or to raise antibodies. They may also
XX be used to screen for agonists or antagonists, which may themselves
XX have use as antibacterial agents. The polynucleotides of the invention
XX may also be used in gene therapy protocols.
XX
XX Sequence 464 AA;
XX
XX Query Match 14.9%; Score 142.5; DB 21; Length 464;
XX Best Local Similarity 23.4%; Pred. No. 8e-07;
XX Matches 45; Conservative 37; Mismatches 71; Indels 39; Gaps 7;
XX
XX 12 KGSLLTLPAYWELSGFHOLTOWAIHYSYKYTWSRF-----KE---LRGKYODGS 60
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 287 KARVKIYTPESLVHGYKVSADKADLFQDVTWTRHSRFDKAEVFEKEKTVKVGK----- 341
XX
XX 61 GYEAFTRKEEYKDNSRFAICTTYSNLDAITLRAGLAYDKAAKST--HLSASIPDTRMWY 118
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 342 -SDRTTTPNWRNTYKVGCGSVQISEPLQLRAGIAFDKSPVRNADYRMNSLPDGNRIWF 400
XX
XX 119 SIGATYKFTPNLSVDVGFALHGRGKKKHVFETO-----NIKGLLLVEADYTTKAT 167
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 401 SAGMKYHIGKNHVVDAAAYTHI-----HINDTYSRTAKASGNDVDSKG-----ASSAREKNH 451
XX
XX 168 ANLYGLNLNRYF 179
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 452 ADIIGLOQTYTKF 463
XX
XX RESULT 7
XX ABB08386
XX ID ABB08386 standard; Protein; 456 AA.
XX AC ABB08386;
XX
XX 18-JUN-2002 (first entry)
XX
XX Tmox amino acid sequence.
XX
XX PHBA; para-Hydrobenzoate; liquid crystal polymer; LCP;
XX toluene monooxygenase; TMO; pcu gene; p-cresol.
XX
XX Pseudomonas mendocina KR-1.
XX
XX WO200192539-A2.
XX

```


CC allows coupling to a carrier in one specific orientation. The
 CC free peptide was used as an immunogen to produce antibodies showing
 CC that the peptide comprises both T-helper determinant and B-cell
 CC epitope(s).
 CC See also AAR12446-R12451 and AAR12453-5 and AAQ12083.

XX Sequence 33 AA;

Query Match 11.6%; Score 111; DB 12; Length 33;
 Best Local Similarity 71.0%; Pred. No. 5,7e-05;
 Matches 22; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 85 LNDALTLRAGLAYDKAASKTHLSASIPDTR 115
 I : |||||:||||:| ||:|||||
 Db 1 LYEKLTLRAGIAYDQAASHRHSRAIPDTR 31

RESULT 10

AAR40062
 ID AAR40062 standard; peptide; 33 AA.

XX AAR40062;

DT 04-FEB-1994 (first entry)

DE H1b OMP P1 peptide H1bP1-10 (339-370).

XX Haemophilus influenzae; type b; H1b; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.

XX Synthetic.

XX Key Location/Qualifiers
 FH MISC-difference 33
 FT /note= "May be absent"

XX W09315205-A.

XX 05-AUG-1993.

XX 03-FEB-1993; 93WO-CA00041.

XX 03-FEB-1992; 92GB-0002219.

XX (CONN-) CONNAUGHT LAB LTD.

PI Chong P, Kandll A, Klein MH, Sia C;

XX WPI; 1993-258681/32.

XX Synthetic Haemophilus influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)

XX Table 1; Page 47; 95pp; English.

XX The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus influenzae type b (H1b) outer membrane proteins P1,
 CC P2 and P6. These peptides may be used in a vaccine against H1b
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise an immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic Pp as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.

XX Sequence 33 AA;

Query Match 11.6%; Score 111; DB 14; Length 33;
 Best Local Similarity 71.0%; Pred. No. 5,7e-05;
 Matches 22; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 85 LNDALTLRAGLAYDKAASKTHLSASIPDTR 115

Db 1 LYEKLTLRAGIAYDQAASHRHSRAIPDTR 31
 I : |||||:||||:| ||:|||||

RESULT 11

AAB46323
 ID AAB46323 standard; Protein; 587 AA.

XX AAB46323;

DT 05-APR-2001 (first entry)

DE H. pylori HPS188 protein.

XX Microbial infection; antibacterial; Helicobacter pylori infection;
 KW vaccine; screening.

XX Helicobacter pylori.

XX W0200073502-A2.

XX 07-DEC-2000.

XX 31-MAY-2000; 2000WO-EP05024.

XX 31-MAY-1999; 99DE-1024965.

XX 17-JUN-1999; 99DE-1027740.

XX 21-JUL-1999; 99DE-1034029.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX (CREA-) CREATOGEN GMBH.

XX Aptel H, Fuchs TM, Gibbs CP, Hueck CJ, Meyer TF;

XX WPI; 2001-049948/06.

XX N-PSDB; AAF25600.

XX Claim 37; Page 274-276; 366pp; German.

XX This invention describes a novel preparation of an agent (A) for
 CC detection, prevention and/or treatment of microbial infection by:
 CC (i) identifying essential genes (I) and corresponding polypeptides
 CC (II); (ii) identifying compounds that are directed against (II) and
 CC inactivate the microbe; (iii) testing these for suitability for use; and
 CC (iv) formulating selected (A). Identifying essential genes (I) comprises
 CC preparation of gene-deficient microorganisms by conditional antisense
 CC inhibition (CAI) and/or subtractive recombination mutagenesis (SRM),
 CC then determining viability and/or survival of the deficient organisms.
 CC The products of the invention have antibacterial activity. (A) (which may
 CC be a nucleic acid (Ia) vector or host cell containing (Ia), derived
 CC polypeptide (Iia), or fragments, (Iia)-specific antibodies or their
 CC fragments or an inhibitor of (Iia)) are particularly used for diagnosis,
 CC treatment or prevention of infection by Helicobacter pylori. Particularly
 CC (Ia) and (Iia) are used in DNA, subunit or live vaccines. The method
 CC identifies essential genes, including those that have homologs in other
 CC species, so identified (A) should have a broad spectrum of activity. Many
 CC gene-deficient cells can be screened quickly, in an automated process,
 CC and the identified genes can be used for screening without purification.

XX Sequence 587 AA;

Query Match 11.3%; Score 108; DB 22; Length 587;
 Best Local Similarity 23.3%; Pred. No. 0.0067;
 Matches 49; Conservative 29; Mismatches 78; Indels 54; Gaps 10;

QY 6 LCPYIG----KGSLLT--KLPAWELSGFHQ-LTDQWAIHYSKYVTEWSRKFEL----- 52
 |||:| ||||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

Db 396 LQPSGLSVLTGKSLNINVSFLQTLSLAYAHQFFKDLHLEGVFTFWSQGNKFLVTPDF 455

QY 93 AGLAYDKAASKTHLSASIPDTRMWSYGATYKFTPNLSVDVGFA 137
 Db 515 GAIDYDQAPSPD-AIGIPDSNGYTVAFGTYNFR---GFDLGVA 555

RESULT 14

AAU27573
 ID AAU27573 standard; Protein: 644 AA.

XX AC AAU27573;
 XX 18-DEC-2001 (first entry)
 XX N. meningitidis fusion protein delta-G287NZ-953.
 DE Neisseria meningitidis fusion protein; ORF46.1;
 KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1;
 KW Neisseria meningitidis.
 XX Synthetic.
 OS Neisseria meningitidis.
 OS Synthetic.
 PN WO200164922-A2.
 XX 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-IB00452.
 XX 28-FEB-2000; 2000GB-0004695.
 PR 13-NOV-2000; 2000GB-0027675.
 XX (CHIR-) CHIRON SPA.
 PA Arico MB, Comanducci M, Galeotti C, Massignani V, Guilianani MM;
 PI Pizza M;
 PI WPI: 2001-582163/65.
 DR N-PSDB; AAS43872.

XX Producing heterologous proteins from Neisseria meningitidis and N.
 PT gonorrhoeae -
 XX
 PS Example 15; Page 41; 119pp; English.
 CC The invention relates to methods for the heterologous expression of
 CC Neisserial proteins from Neisseria meningitidis and Neisseria
 CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the
 CC leader peptide, and may be replaced by a domain from a different protein
 CC to make a fusion protein, in order to enhance heterologous expression of
 CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine
 CC stretch, can be mutated to enhance expression. The proteins used in the
 CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
 CC AAU27553-AAU27610 represent Neisserial proteins and peptide regions of
 CC proteins of the invention.
 XX SQ Sequence 644 AA;

Query Match 10.2%; Score 98; DB 22; Length 644;
 Best Local Similarity 29.7%; Pred. No. 0.094;
 Matches 51; Conservative 24; Mismatches 69; Indels 28; Gaps 11;

QY 3 NGVLGPYI--CKGSITLKL--PAYWELSGFHQLTDQWAIHYSYKYTEWSR--FKELRG-K 55

Db 418 NGFGTWTENGDDVSGKFGYCPAGEEVAG-----KYSYRPTDAEKGFGVFGAKK 467

QY 56 YQDGSQGYEAF--KKEEYKDNRSRFAI---GTTYSLNDALTLAGLAYDKAASKTHLSASIP 111

Db 468 EQDGSQGGGATYKVDEYHANAFAIDHFTSTNVGGFVGLTGSVEFDQAKRDGKIDITIP 527

QY 112 DTRMWSYGATYKFTPNL-SVDVGFALHRCCKKHVETQ-NIKGLLLVEAD 161

Db 528 VANLQ-----SGSQHFTDHLKSADI-FDAAQYPDIRFVSTKFNFGKLLVSD 574

RESULT 15

AAE10018
 ID AAE10018 standard; Protein: 644 AA.

XX AC AAE10018;
 XX 29-NOV-2001 (first entry)
 XX N. meningitidis strain 394/98 ORF46.1 delta G287NZ-953 fusion protein.
 DE Heterologous expression; Neisserial protein; open reading frame; ORF;
 KW ORF46.1 delta G287NZ-953 fusion protein.
 XX Neisseria meningitidis.
 XX WO200164920-A2.
 PN 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-IB00420.
 XX 28-FEB-2000; 2000GB-0004695.
 PR 13-NOV-2000; 2000GB-0027675.
 XX (CHIR-) CHIRON SPA.
 PA Arico MB, Comanducci M, Galeotti C, Massignani V, Guilianani MM;
 PI Pizza M;
 PI WPI: 2001-557776/62.
 DR N-PSDB; AADI7037.

XX Heterologous expression for the expression of two or more Neisserial
 PT proteins in fused state -
 XX
 PS Example 2; Page 11-12; 52pp; English.
 CC The present invention relates to a method for simultaneous heterologous
 CC expression of two or more Neisserial proteins which are in a fused
 CC state. The method is useful for simultaneous heterologous expression of
 CC two or more Neisserial proteins. A protein that may be unstable or
 CC poorly expressed on its own is assisted by adding a suitable hybrid
 CC partner and commercial manufacture is simplified-only one expression and
 CC purification need to be employed in order to produce two separately-
 CC useful proteins. The present sequence is Neisseria meningitidis
 CC (serogroup B, New Zealand strain 394/98) ORF46.1 (open reading frame)
 CC delta G287NZ-953 fusion.
 XX SQ Sequence 644 AA;

Query Match 10.2%; Score 98; DB 22; Length 644;
 Best Local Similarity 29.7%; Pred. No. 0.094;
 Matches 51; Conservative 24; Mismatches 69; Indels 28; Gaps 11;

QY 3 NGVLGPYI--CKGSITLKL--PAYWELSGFHQLTDQWAIHYSYKYTEWSR--FKELRG-K 55

Db 418 NGFGTWTENGDDVSGKFGYCPAGEEVAG-----KYSYRPTDAEKGFGVFGAKK 467

QY 56 YQDGSQGYEAF--KKEEYKDNRSRFAI---GTTYSLNDALTLAGLAYDKAASKTHLSASIP 111

Db 468 EQDGSQGGGATYKVDEYHANAFAIDHFTSTNVGGFVGLTGSVEFDQAKRDGKIDITIP 527

QY 112 DTRMWSYGATYKFTPNL-SVDVGFALHRCCKKHVETQ-NIKGLLLVEAD 161

Db 528 VANLQ-----SGSQHFTDHLKSADI-FDAAQYPDIRFVSTKFNFGKLLVSD 574

RESULT 16

AAU75566
 ID AAU75566 standard; Protein: 708 AA.

XX AC AAU75566;
 XX


```

XX PR 11-APR-2000; 2000FR-0004629.
XX PA (INSP ) INST PASTEUR.
XX PI Buchrleser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussurget O, Chetouani F, Nedjarl H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;
XX DR WPI; 2002-010914/01.
XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and
XX PT related polypeptides .
XX PS Claim 6; SEQ ID No 32; 192pp; French.
XX CC The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded by the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 572 AA;

Query Match 9.8%; Score 94; DB 23; Length 572;
Best Local Similarity 26.1%; Pred. No. 0.22;
Matches 57; Conservative 30; Mismatches 67; Indels 64; Gaps 12;

QY 4 GVLGPYIGKSLTLKLPAYWELSGFHLTDQWAIHYKYKTEWSRPFKELR-GKYQDGSY 62
DB 140 GIGSYNGK-SVTMKTWEYSDSGWYQINAFKYPKSHKESLEDNAKLRNGPSWDSYY 198
QY 63 EAFTKK--EYKQ-----NSRFAIGTYS--LN-----D 87
DB 199 KGARENAKYKDATANLOGRYATDNTYASKLNTLISSYNTQYDTLYDTIKQKNVSD 258
QY 88 ALTLRA-----GLAYDKAAKTHLSASIPDTR-----MMYSIGATYKFTPNLSVD 133
DB 259 AKVVKADGHHGVSYGIYNTSAASAKLSTGAPYNNKVKILKEGTTSRGTWVQFSLNNKV- 317
QY 134 VGFAHLRGKKKHVF---ETONIKGLLLVEADYTTTKATA 168
DB 318 IGWM---DKRAFYYPKATNVKNTLNL-----TGKITA 346

RESULT 21
AAB07698
ID AAB07698 standard; Protein; 691 AA.
XX AC AAB07698;
XX CC
XX DT 07-NOV-2000 (first entry)
XX DE A Neisseria meningitidis BASB053 polypeptide.

```

```

XX BASB053; Neisseria meningitidis infection; vaccine.
XX OS Neisseria meningitidis.
XX PN W0200042193-AL.
XX PD 20-JUL-2000.
XX PF 10-JAN-2000; 2000WO-EP00137.
XX PR 15-JAN-1999; 99GB-0000959.
XX PR 28-JAN-1999; 99GB-0001903.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Ruelle J;
XX DR WPI; 2000-476062/41.
XX DR N-PSDB; AAA59217.
XX PT New Neisseria meningitidis polypeptide useful for diagnosis of
XX PT Neisseria infection and for development of vaccines against such
XX PT infection .
XX PS Claim 3; Page 56; 92pp; English.
XX CC The present sequence represents a Neisseria meningitidis BASB053
XX CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
XX CC for BASB053 may be identified in a biological sample in order to
XX CC diagnose a Neisseria meningitidis infection in an animal. The BASB053
XX CC polypeptides and polynucleotides may be used as vaccines, for
XX CC generating an immune response in an animal. A composition comprising
XX CC at least one antibody immunospecific for BASB053 may be used to
XX CC treat humans infected with Neisseria meningitidis.
XX SQ Sequence 691 AA;

Query Match 9.8%; Score 93.5; DB 21; Length 691;
Best Local Similarity 23.5%; Pred. No. 0.32;
Matches 43; Conservative 23; Mismatches 70; Indels 47; Gaps 9;

QY 6 LGPYIGKSLTLKLPAYWELSGFHLTDQWAIHYKYKTEWSRPFKELRGKYQDGSYDAF 65
DB 547 LGKRVMEGVET-----EISG--AVTPKQIHAGYSYLH-SQIKTASNSRDDG---IF 592
QY 66 TKREYKDNRSFAIGTYSYSLNDALTLRAGLAYDKAAKTHLSASIPDTRMMYSIGATYK 125
DB 593 LLMPKHSAN-----LWTTYQVTPELTIGGV---NAMSGITSSAGMIAGGYATFDAMAAYR 645
QY 126 FTPNLSVDVGFALHGRKKKHEVETONIKGLLLVEADYTTKATANYGL-----NLN 176
DB 646 FTPKLLQINADNI--FNRYH-----YARVGGANTFNPGSERSLTANLR 688
QY 177 YRF 179
DB 689 YSF 691

RESULT 22
AAM50721
ID AAM50721 standard; Protein; 344 AA.
XX AC AAM50721;
XX DT 18-APR-2002 (first entry)
XX DE Haemophilus paragallinarum strain 0222 haemagglutinin.
XX KW Haemagglutinin; HgaA; antigen; vaccine; immunisation; coryza;
XX KW chicken.
XX OS Haemophilus paragallinarum.

```

```

XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /label= Signal_peptide
XX FT Protein 22..344
XX FT /label= Mature_protein
XX PN WO200204485-A1.
XX PD 17-JAN-2002.
XX PF 06-JUL-2001; 2001WO-AU00822.
XX PR 07-JUL-2000; 2000AU-0008652.
XX PA (UYOU ) UNIV QUEENSLAND.
XX PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX DR WPI; 2002-154917/20.
XX DR N-PSDB; ABA91425.
XX PT New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
XX PT useful as a vaccine for immunising chickens against coryza caused by
XX PT the species
XX PS Claim 2; Flg 4; 67pp; English.
XX CC The present sequence is that of the HagA haemagglutinin of
XX CC Haemophilus paragallinarum strain 0222 (serovar B), the causative
XX CC agent of infectious coryza in chickens. The invention provides
XX CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
XX CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
XX CC A, B and C) of H. paragallinarum, the causative agent of infectious
XX CC coryza of chickens. The polypeptides are useful in vaccines for
XX CC immunisation against infectious coryza, as are the nucleic acids
XX CC when expressed in attenuated bacteria, especially Salmonella or
XX CC Mycoplasma (claimed). The recombinant polypeptide is preferably
XX CC the mature protein, or a biologically active fragment, variant or
XX CC derivative, that is capable of eliciting an immune response,
XX CC providing protection against one or more strains of H.
XX CC paragallinarum in chickens. Also claimed are methods of using the
XX CC haemagglutinin polypeptides and nucleic acids for detection and
XX CC diagnosis of infectious coryza in chickens.
XX SQ Sequence 344 AA;

Query Match 9.7%; Score 93; DB 23; Length 344;
Best Local Similarity 20.7%; Pred. No. 0.14;
Matches 38; Conservative 31; Mismatches 71; Indels 44; Gaps 7;

QY .16 TLKLPAYWELSGFHOLTDOMAIHYSY-----KYTEWSRFKELRGKYODG 59
DB 55 TLRNSVTYGVGGYQITDNFAVELGYDDFGRAKRQGGTIVIKYTNHGAHLSLKASYPVL 114
QY 60 SCYEAFTR-----KEEYKDNSRFAIGTT--YSLNDALTLRAGLAYDKAASKTHLSASIP 111
DB 115 EGLDVIARVGAALIRSDYKPTKRAAPNQTHEHSLKVSVPFAGGLEYN-----LPSLP 166
QY 112 D-----TDRMWS-----IGATYKFTPNL-SVDVGFAGLRGKKHFFVETONIKGLLVE 159
DB 167 ELALRVEYQWNVKVRVEKDGSRVDYTPSIGSVTAGLSYRFQOSAPVPEPKVAKTFALN 226
QY 160 ADYT 163
DB 227 SDVT 230

RESULT 23
AAM50724
ID AAM50724 standard; Protein: 344 AA.
XX AC AAM50724;

```

```

XX DT 18-APR-2002 (first entry)
XX DE Haemophilus paragallinarum strain H-18 haemagglutinin.
XX KW Haemagglutinin; HagA; antigen; vaccine; immunisation; coryza;
XX KW chicken.
XX OS Haemophilus paragallinarum.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /label= Signal_peptide
XX FT Protein 22..344
XX FT /label= Mature_protein
XX PN WO200204485-A1.
XX PD 17-JAN-2002.
XX PF 06-JUL-2001; 2001WO-AU00822.
XX PR 07-JUL-2000; 2000AU-0008652.
XX PA (UYOU ) UNIV QUEENSLAND.
XX PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX DR WPI; 2002-154917/20.
XX DR N-PSDB; ABA91425.
XX PT New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
XX PT useful as a vaccine for immunising chickens against coryza caused by
XX PT the species
XX PS Claim 2; Fig 4; 67pp; English.
XX CC The present sequence is that of the HagA haemagglutinin of
XX CC Haemophilus paragallinarum strain H-18 (serovar C), the causative
XX CC agent of infectious coryza in chickens. The invention provides
XX CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
XX CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
XX CC A, B and C) of H. paragallinarum, the causative agent of infectious
XX CC coryza of chickens. The polypeptides are useful in vaccines for
XX CC immunisation against infectious coryza, as are the nucleic acids
XX CC when expressed in attenuated bacteria, especially Salmonella or
XX CC Mycoplasma (claimed). The recombinant polypeptide is preferably
XX CC the mature protein, or a biologically active fragment, variant or
XX CC derivative, that is capable of eliciting an immune response,
XX CC providing protection against one or more strains of H.
XX CC paragallinarum in chickens. Also claimed are methods of using the
XX CC haemagglutinin polypeptides and nucleic acids for detection and
XX CC diagnosis of infectious coryza in chickens.
XX SQ Sequence 344 AA;

Query Match 9.7%; Score 93; DB 23; Length 344;
Best Local Similarity 20.7%; Pred. No. 0.14;
Matches 38; Conservative 31; Mismatches 71; Indels 44; Gaps 7;

QY 16 TLKLPAYWELSGFHOLTDOMAIHYSY-----KYTEWSRFKELRGKYODG 59
DB 55 TLRNSVTYGVGGYQITDNFAVELGYDDFGRAKRQGGTIVIKYTNHGAHLSLKASYPVL 114
QY 60 SCYEAFTR-----KEEYKDNSRFAIGTT--YSLNDALTLRAGLAYDKAASKTHLSASIP 111
DB 115 EGLDVIARVGAALIRSDYKPTKRAAPNQTHEHSLKVSVPFAGGLEYN-----LPSLP 166
QY 112 D-----TDRMWS-----IGATYKFTPNL-SVDVGFAGLRGKKHFFVETONIKGLLVE 159
DB 167 ELALRVEYQWNVKVRVEKDGSRVDYTPSIGSVTAGLSYRFQOSAPVPEPKVAKTFALN 226
QY 160 ADYT 163

```


Matches	37;	Conservative	30;	Mismatches	73;	Indels	44;	Gaps	67;
Qy	16	TLKLPAYWELSGFHQLTDOWAIHYSY-----KYTESRRFKELRGKYODG	59						
Dd	55	TLRNSVTYGVEGGQITDNFAFALCYDDFGRAKLQRDGETVGKHTNHGAHLSLKASYPVL	114						
Qy	60	SGYEAFYK-----KEEYKDNSRFAIGTT--YSLNDALTLAGLAYDKAASKTHLSASIP	111						
Dd	115	EGLDVAVRVGAALIRSDYKPTKRAAPNETHSHLSLKVPVFAGGLEYN-----LPSLP	166						
Qy	112	DTD-----RMWYSICATYKFTPNL-SVDVGFPAHLGRKKKHFFVETONIKGLLIVE	159						
Dd	167	ELAURVEYQWNVKVRWEKGDSRDYDTPSIGSVTAGUSYRFQSQSAVPVEPKVAKTFAIN	226						
Qy	160	ADYT 163							
Dd	227	SDVT 230							
RESULT 26									
AAM50717	ID	AAM50717 standard; Protein; 344 AA.							
XX	AC	AAM50717;							
XX	AC	AAM50717;							
DT	DT	18-APR-2002 (first entry)							
DE	DE	Haemophilus paragallinarum strain 221 haemagglutinin.							
KW	KW	Haemagglutinin; HaggA; antigen; vaccine; immunisation; coryza; chicken.							
KW	KW	Haemophilus paragallinarum.							
OS	OS								
FH	FH	Key Location/Qualifiers							
FT	FT	Peptide 1..21							
FT	FT	/label= Signal_peptide							
FT	FT	22..344							
FT	FT	/label= Mature_protein							
XX	XX								
PN	PN	WO200204485-A1.							
XX	XX								
PD	PD	17-JAN-2002.							
XX	XX								
PF	PF	06-JUL-2001; 2001WO-AU00822.							
XX	XX								
PR	PR	07-JUL-2000; 2000AU-0008652.							
XX	XX	(UYQU) UNIV QUEENSLAND.							
PA	PA								
XX	XX								
PI	PI	Terry TD, Tseng H, Hobb RI, Jennings MP; Downes J;							
XX	XX								
DR	DR	WPI: 2002-154917/20.							
DR	DR	N-PSDB; ABA91418.							
XX	XX								
PT	PT	New haemagglutinin polypeptide isolated from Haemophilus paragallinarum							
PT	PT	useful as a vaccine for immunising chickens against coryza caused by							
PT	PT	the species							
XX	XX								
PS	PS	Claim 2; Fig 4; 67pp; English.							
XX	XX								
CC	CC	The present sequence is that of the HaggA haemagglutinin of							
CC	CC	Haemophilus paragallinarum strain 221 (serovar A), the causative							
CC	CC	agent of infectious coryza in chickens. The invention provides							
CC	CC	recombinant haemagglutinin polypeptides (see AAM50716-27) and							
CC	CC	encoding nucleic acids (see ABA91417-27) from 11 strains (serovars							
CC	CC	A, B and C) of H. paragallinarum, the causative agent of infectious							
CC	CC	coryza of chickens. The polypeptides are useful in vaccines for							
CC	CC	immunisation against infectious coryza, as are the nucleic acids							
CC	CC	when expressed in attenuated bacteria, especially Salmonella or							
CC	CC	Mycoplasma (claimed). The recombinant polypeptide is preferably							
CC	CC	the mature protein, or a biologically active fragment, variant or							
CC	CC	derivative, that is capable of eliciting an immune response,							


```
PR 07-JUL-2000; 2000AU-0008652.
PA (UYOU ) UNIV QUEENSLAND.
PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX WPI; 2002-154917/20.
XX N-PSDB; ABA91424.
DR New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
PT useful as a vaccine for immunising chickens against coryza caused by
PT the species
XX Claim 2; Fig 4; 67pp; English.
XX The present sequence is that of the HagA haemagglutinin of
CC Haemophilus paragallinarum strain Modesto (serovar C), the causative
CC agent of infectious coryza in chickens. The invention provides
CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
CC A, B and C) of H. paragallinarum, the causative agent of infectious
CC coryza of chickens. The polypeptides are useful in vaccines for
CC immunisation against infectious coryza, as are the nucleic acids
CC when expressed in attenuated bacteria, especially Salmonella or
CC Mycoplasma (claimed). The recombinant polypeptide is preferably
CC the mature protein, or a biologically active fragment, variant or
CC derivative, that is capable of eliciting an immune response,
CC providing protection against one or more strains of H.
CC paragallinarum in chickens. Also claimed are methods of using the
CC haemagglutinin polypeptides and nucleic acids for detection and
CC diagnosis of infectious coryza in chickens.
XX Sequence 341 AA;
Query Match 9.4%; Score 89.5; DB 23; Length 341;
Best Local Similarity 20.4%; Pred. No. 0.33;
Matches 37; Conservative 32; Mismatches 71; Indels 41; Gaps 7;
QY 16 TLKLPAYWELSGFHQLTDQWAIHYSY-----KYTEWSRFRKELRGKYQDG 59
DB 55 TLRNSVTYGVFGGYQITDNFAVELGYDDFGRAKFRQDGETVIKHTNHGAHLSKASYPVL 114
QY 60 SGYEAFK-----KEEYKDNSRFAIGTT--YSLNDALTLRAGLAYDKAASKTHLSASIP 111
DB 115 EGLDVIYARVGAALIRSDYKPTKRAAPNETHEHSLKVPFAGGLEYN-----LPSLP 166
QY 112 D-----TDRMWS-----IGATYKFTPNL--SVDVGFALRGKKHFEVETONIKGLLLEADY 162
DB 167 ELALRVEYQWNVKVGDRGSRVDYTPSIGSVTAGLSYRFQGSAPVVEPKVAKTFALNSDV 226
QY 163 T 163
DB 227 T 227
RESULT 30
AAM50726
ID AAM50726 standard; Protein; 341 AA.
XX AC AAM50726;
XX DT 18-APR-2002 (first entry)
XX DE Haemophilus paragallinarum strain HP60 haemagglutinin.
KW Haemagglutinin; HagA; antigen; vaccine; immunisation; coryza;
KW chicken.
XX OS Haemophilus paragallinarum.
XX FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= Signal_peptide
```

```
FT Protein 22..341
XX /label= Mature_protein
PN WO200204485-A1.
XX PD 17-JAN-2002.
XX PF 06-JUL-2001; 2001WO-AU00822.
XX PR 07-JUL-2000; 2000AU-0008652.
XX PA (UYOU ) UNIV QUEENSLAND.
XX PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX WPI; 2002-154917/20.
XX N-PSDB; ABA91427.
PT New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
PT useful as a vaccine for immunising chickens against coryza caused by
PT the species
XX Claim 2; Fig 4; 67pp; English.
XX The present sequence is that of the HagA haemagglutinin of
CC Haemophilus paragallinarum strain HP60 (serovar C), the causative
CC agent of infectious coryza in chickens. The invention provides
CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
CC A, B and C) of H. paragallinarum, the causative agent of infectious
CC coryza of chickens. The polypeptides are useful in vaccines for
CC immunisation against infectious coryza, as are the nucleic acids
CC when expressed in attenuated bacteria, especially Salmonella or
CC Mycoplasma (claimed). The recombinant polypeptide is preferably
CC the mature protein, or a biologically active fragment, variant or
CC derivative, that is capable of eliciting an immune response,
CC providing protection against one or more strains of H.
CC paragallinarum in chickens. Also claimed are methods of using the
CC haemagglutinin polypeptides and nucleic acids for detection and
CC diagnosis of infectious coryza in chickens.
XX Sequence 341 AA;
Query Match 9.4%; Score 89.5; DB 23; Length 341;
Best Local Similarity 20.4%; Pred. No. 0.33;
Matches 37; Conservative 32; Mismatches 71; Indels 41; Gaps 7;
QY 16 TLKLPAYWELSGFHQLTDQWAIHYSY-----KYTEWSRFRKELRGKYQDG 59
DB 55 TLRNSVTYGVFGGYQITDNFAVELGYDDFGRAKFRQDGETVIKHTNHGAHLSKASYPVL 114
QY 60 SGYEAFK-----KEEYKDNSRFAIGTT--YSLNDALTLRAGLAYDKAASKTHLSASIP 111
DB 115 EGLDVIYARVGAALIRSDYKPTKRAAPNETHEHSLKVPFAGGLEYN-----LPSLP 166
QY 112 D-----TDRMWS-----IGATYKFTPNL--SVDVGFALRGKKHFEVETONIKGLLLEADY 162
DB 167 ELALRVEYQWNVKVGDRGSRVDYTPSIGSVTAGLSYRFQGSAPVVEPKVAKTFALNSDV 226
QY 163 T 163
DB 227 T 227
RESULT 31
AAR77432
ID AAR77432 standard; Protein; 459 AA.
XX AC AAR77432;
XX DT 22-JUN-1996 (first entry)
XX DE Moraxella catarrhalis outer membrane protein E.
```

XX Bacterium; Branhamella catarrhalis; sinusitis; conjunctivitis;
KW pneumonia; endocarditis; septicaemia; meningitis; otitis media;
KW lower respiratory tract infection; chronic bronchitis;
KW chronic obstructive pulmonary disease; vaccine; diagnostic;
KW immunoassay.
XX
OS Moraxella catarrhalis.
XX
FH Key Location/Qualifiers
FT Peptide 1..25
FT Region /label= sig_peptide 23..34
FT Region /note= "antigenic determinant hydrophilic region" 234..299
FT Region /note= "antigenic determinant hydrophilic region" 369..374
FT Region /note= "antigenic determinant hydrophilic region"
XX
FN W09531215-AL.
XX
PD 23-NOV-1995.
XX
PF 20-APR-1995; 95WO-US05134.
XX
PR 17-MAY-1994; 94US-0245758.
XX
PA (UUNY) UNIV NEW YORK STATE RES FOUND.
XX
PI Bhushan R, Murphy TF;
XX
DR WPI; 1996-010692/01.
DR N-PSDB; AAT04376.
XX
PT Vaccine contg outer membrane protein E of Moraxella catarrhalis -
PT for the detection of M.catarrhalis-specific antisera and in
PT diagnostic immunoassays
XX
PS Claim 11; Page 40-43; 58pp; English.
XX
CC This E protein from M. catarrhalis strain 25240 may be produced
CC recombinantly from cells cultured from a host cell system engineered
CC to include a vector containing a nucleotide sequence that regulates
CC expression of DNA encoding E epitopes. The host cell system may be
CC bacteria, yeast, filamentous fungi, insect and mammalian cell lines.
CC The protein may be used to detect M. catarrhalis-specific antisera
CC in body fluid and may be used in a vaccine against M. catarrhalis.
CC It may also be used as an antigen in diagnostic immunoassays.
XX
SQ Sequence 459 AA;

Query Match 9.3%; Score 89; DB 17; Length 459;
Best Local Similarity 23.2%; Pred. No. 0.57;
Matches 42; Conservative 19; Mismatches 84; Indels 36; Gaps 6;

QY 9 YIGKSGTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKEKRYQDGSYEAFKK 68
DB 292 YAGKDKFKVTLPSWNLDFQTVGNPTLLTAKVRYVPWSDFDIRPTQYTE-----TTK 344

QY 69 EEEY-----KDNSRFAIGTYSINDALTLAGLAYDKAASKTHLSASIPDTRMWY 118
DB 345 LRYPOGLPIISYDKDQWSAEVGLGKRVSDRLAVSGAVGWSGAGN---PASSLGPIKGY 401

QY 119 S---IGATYKFTPNLSVDVGFALHGRKKHIFVETQNIKGLLLVEADYTTKATANLYGLNLN 176
DB 402 SLGLGARYNVTPEWSLSLG-----GKYFKEGDAQ-----AQLPTKDKVGNFDSNDG 447

QY 177 Y 177
DB 448 Y 448

RESULT 32

AAAY311154
ID AAAY311154 standard; Protein; 459 AA.
AC AAAY311154;
XX
DT 26-OCT-1999 (first entry)
XX
DE M. catarrhalis outer membrane E protein.
XX
KW Outer membrane protein; E protein; vaccine; antigen; epitope; therapy;
KW heat-modifiable protein; immunogen; prophylactic; active immunization;
KW antisera; passive immunization; prevention; otitis media; sinusitis;
KW conjunctivitis; lower respiratory tract infection; diagnosis;
KW immunoassay.
XX
OS Moraxella catarrhalis.
XX
FH Key Location/Qualifiers
FT Peptide 1..25
FT Protein /label= signal_peptide 26..459
XX
PN US5948412-A.
XX
PD 07-SEP-1999.
XX
PF 03-MAR-1997; 97US-0810655.
XX
PR 03-MAR-1997; 97US-0810655.
PR 17-MAY-1994; 94US-0245758.
XX
PA (UUNY) UNIV NEW YORK STATE RES FOUND.
XX
PI Murphy TF;
XX
DR WPI; 1999-S17930/43.
DR N-PSDB; AA209296.
XX
PT Antigenic peptide, oligopeptide and protein useful as vaccine for
PT Moraxella catarrhalis
XX
PS Claim 1; Column 25-28; 20pp; English.
XX
CC This invention describes a novel pure antigenic peptide, oligopeptide or
CC protein (I) with one or more epitopes of E, an outer membrane protein of
CC Moraxella catarrhalis. The E protein of the invention has an apparent
CC molecular weight of 35000-50000 daltons by sodium dodecyl
CC sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and is a
CC heat-modifiable protein. (I) can be used as immunogens in prophylactic
CC and/or therapeutic vaccine formulations for active immunization and for
CC generating protein-specific and peptide-specific antisera useful for
CC passive immunization. Antigenic formulations comprising (I) can be used
CC to prevent otitis media, sinusitis, conjunctivitis and lower respiratory
CC tract infections caused by Moraxella catarrhalis. (I) can be used as
CC antigens for diagnostic immunoassays. This sequence represents the
CC Moraxella catarrhalis strain 25240 outer membrane E protein which is
CC described in the method of the invention.
XX
SQ Sequence 459 AA;

Query Match 9.3%; Score 89; DB 20; Length 459;
Best Local Similarity 23.2%; Pred. No. 0.57;
Matches 42; Conservative 19; Mismatches 84; Indels 36; Gaps 6;

QY 9 YIGKSGTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKEKRYQDGSYEAFKK 68
DB 292 YAGKDKFKVTLPSWNLDFQTVGNPTLLTAKVRYVPWSDFDIRPTQYTE-----TTK 344

QY 69 EEEY-----KDNSRFAIGTYSINDALTLAGLAYDKAASKTHLSASIPDTRMWY 118
DB 345 LRYPOGLPIISYDKDQWSAEVGLGKRVSDRLAVSGAVGWSGAGN---PASSLGPIKGY 401

QY 119 S---IGATYKFTPNLSVDVGFALHGRKKHIFVETQNIKGLLLVEADYTTKATANLYGLNLN 176

Db 402 SLGLGARYNVTPWLSLGL-----GKYKFGDAQ-----AQLPTKDKVGNFDSNDG 447
 Qy 177 Y 177
 Db 448 Y 448

RESULT 33
 AAM50718
 ID AAM50718 standard; Protein; 344 AA.
 AC AAM50718;
 XX
 DT 18-APR-2002 (first entry)
 DE Haemophilus paragallinarum strain 2403 haemagglutinin.
 KW Haemagglutinin; HAgA; antigen; vaccine; immunisation; coryza;
 KW chicken.
 XX
 OS Haemophilus paragallinarum.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 FT Protein 22..344
 FT /label= Mature_protein
 XX
 PN WO200204485-A1.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-AU00822.
 XX
 PR 07-JUL-2000; 2000AU-0008652.
 XX
 PA (UYQU) UNIV QUEENSLAND.
 XX
 PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 XX
 DR WPI; 2002-154917/20.
 DR N-PSDB; ABA91419.
 XX
 PT New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species
 XX
 PS Claim 2; Fig 4; 67pp; English.
 XX
 CC The present sequence is that of the HAgA haemagglutinin of
 CC Haemophilus paragallinarum strain 2403 (serovar A), the causative
 CC agent of infectious coryza in chickens. The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of H. paragallinarum, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially Salmonella or
 CC Mycoplasma (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response.
 CC providing protection against one or more strains of H.
 CC paragallinarum in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.
 XX
 SQ Sequence 344 AA;

Query Match 9.2%; Score 88; DB 23; Length 344;
 Best Local Similarity 20.1%; Pred. No. 0.49;
 Matches 37; Conservative 32; Mismatches 71; Indels 44; Gaps 7;

Qy 16 TLKLPAYWELSGFHLQTDQWAIHYSY-----KYTEWSRFKELRGKYODG 59
 Db 55 TLRNVTYGVFGGQITDNFAVELGYDDFGRAKSRQGGETVIKHTNHGAHLSLRASYPVL 114
 Qy 60 SGYEAFYK-----KEEYKDNSRFAIGTT--YSLNDALTLAGLAYDKAASKTHLSASIP 111
 Db 115 EGLDYARVGVGAALIRSDYKPTKRAAPNOTHEHSLKVSFVAGGLEYN-----LPSLP 166
 Qy 112 D---TDRMWSY-----IGATYKFTPNL--SDVGVFAHLRGKKKHVFETQNIKGLLIVE 159
 Db 167 ELALRVEYQWNVKVRVEKSGSRVDYTPSGISVTAGLSYREGQSAPVVEPKVWAKTFALN 226
 Qy 160 ADYT 163
 Db 227 SDVT 230

RESULT 34
 AAM50722
 ID AAM50722 standard; Protein; 344 AA.
 AC AAM50722;
 XX
 DT 18-APR-2002 (first entry)
 DE Haemophilus paragallinarum strain 2671 haemagglutinin.
 KW Haemagglutinin; HAgA; antigen; vaccine; immunisation; coryza;
 KW chicken.
 XX
 OS Haemophilus paragallinarum.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 FT Protein 22..344
 FT /label= Mature_protein
 XX
 PN WO200204485-A1.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-AU00822.
 XX
 PR 07-JUL-2000; 2000AU-0008652.
 XX
 PA (UYQU) UNIV QUEENSLAND.
 XX
 PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 XX
 DR WPI; 2002-154917/20.
 DR N-PSDB; ABA91423.
 XX
 PT New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species
 XX
 PS Claim 2; Fig 4; 67pp; English.
 XX
 CC The present sequence is that of the HAgA haemagglutinin of
 CC Haemophilus paragallinarum strain 2671 (serovar B), the causative
 CC agent of infectious coryza in chickens. The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of H. paragallinarum, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially Salmonella or
 CC Mycoplasma (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response.
 CC providing protection against one or more strains of H.
 CC paragallinarum in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.

CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.

SQ Sequence 344 AA;

Query Match 9.2%; Score 88; DB 23; Length 344;
 Best Local Similarity 20.1%; Pred. No. 0.49;
 Matches 37; Conservative 32; Mismatches 71; Indels 44; Gaps 7;

QY 16 TLKLPAYWELSGPHOLTDOAHISY-----KYTEWSRFEKELRGKYQDG 59
 DB 55 TLRNVTYGVGGYQITDNFAVELGYDDFGRAKRROGGETVIRKHTNHGAHLSLKASYPVL 114
 QY 60 SGYEAFK-----KEEYKDNSFAIGTT--YSLNDALTLRAGLAYDKAAASKTHLSASIP 111
 DB 115 EGLDVARVGAALIRSDYKPTKRAAPNQTHEHSLSKVPFAGGLEYN-----LPSLP 166
 QY 112 D-----TDRMYS-----IGATYKFTPNL-SVDVGFAHLRGKKHFEVETQNIKGLLVE 159
 DB 167 ELALRVEYQMVNKGVRVEKDGSRVDYTPSIGSVTAGLSYRFGSAPVVEPKVAKTFALN 226
 QY 160 ADVT 163
 DB 227 SDVT 230

RESULT 35

AAR12455
 ID AAR12455 standard; Protein; 36 AA.

AC AAR12455;

DT 26-JUL-1991 (first entry)

DE HIBPI-14(400-437) 6U H.influenzae Type b OMP epitope.

XX bacterial meningitis; vaccine; PI gene; T-cell antigen;
 KW outer membrane protein.

XX Haemophilus influenzae.

XX WO9106652-A.

XX 16-MAY-1991.

XX 31-OCT-1990; 90WO-CA00374.

XX 31-OCT-1989; 89GB-0024473.

XX (CONN-) CONNAUGHT LAB LTD.
 PA (UNIW) WASHINGTON UNIV ST LOUIS.

XX Munson RS, Grass S, Chong P, Yang Y, Fahim R, McVerry P;
 PI Klein M;

XX WPI; 1991-164201/22.

XX Outer membrane protein of Haemophilus influenzae type B - used as
 PT vaccine against infections, esp. in infants and for diagnosis

XX Disclosure; Table 2; 33pp; English.

XX Synthetic peptide corresponding to amino acids 400 to 437 of the H.
 CC Influenzae strain 8358 (OMP subtype 6U) PI protein. The N-terminal
 CC Cys residue is additional and allows coupling to a carrier in one
 CC specific orientation. The free peptide was used as an immunogen to
 CC produce antibodies showing that the peptide comprises both T-helper
 CC determinant and B-cell epitope(s).
 CC See also AAR12446-R12454 and AAQ12083.

XX Sequence 36 AA;

Query Match 9.1%; Score 87.5; DB 12; Length 36;

Best Local Similarity 54.3%; Pred. No. 0.024;
 Matches 19; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
 QY 146 FVETQNIK-GLLVEADYTTKATANLYGLNLNRYF 179
 DB 2 FKEAQAAGGFITTTANTYTSQAHANLYGLNLNYSF 36

RESULT 36

AAR40066
 ID AAR40066 standard; peptide; 36 AA.

AC AAR40066;

XX 04-FEB-1994 (first entry)

XX H1b-OMP PI peptide HIBPI-14 (400-437).

XX Haemophilus influenzae; type b; H1b; outer membrane protein; PI; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.

XX Synthetic.

XX Key Location/Qualifiers
 FT Misc-difference 1
 FT /note- "May be absent"

FT Misc-difference 5..14
 FT /note- "Residues are different to those found in PI
 FT protein of H. influenzae strain 1H"

XX WO9315205-A.

XX 05-AUG-1993.

XX 03-FEB-1993; 93WO-CA00041.

XX 03-FEB-1992; 92GB-0002219.

XX (CONN-) CONNAUGHT LAB LTD.

XX Chong P, Kandil A, Klein MH, Sia C;

XX WPI; 1993-258681/32.

XX Synthetic Haemophilus influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)

XX Table 1; Page 48; 99pp; English.

XX The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus influenzae type b (H1b) outer membrane proteins PI,
 CC P2 and P6. These peptides may be used in a vaccine against H1b
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise an immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.

XX Sequence 36 AA;

Query Match 9.1%; Score 87.5; DB 14; Length 36;
 Best Local Similarity 54.3%; Pred. No. 0.024;
 Matches 19; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 146 FVETQNIK-GLLVEADYTTKATANLYGLNLNRYF 179
 DB 2 FKEAQAAGGFITTTANTYTSQAHANLYGLNLNYSF 36

RESULT 37

AAW50725

PS Claim 2; Fig 4; 67pp; English.

XX The present sequence is a consensus of the HaqA haemagglutinins
CC of 11 strains of Haemophilus paragallinarum serovars A, B and C
CC (see AAM50716-26). The invention provides recombinant haemagglutinin
CC polypeptides and encoding nucleic acids (see ABA91417-27) from H.
CC paragallinarum, the causative agent of infectious coryza of chickens.
CC The polypeptides are useful in vaccines for immunisation against
CC infectious coryza, as are the nucleic acids when expressed in
CC attenuated bacteria, especially Salmonella or Mycoplasma (claimed).
CC The recombinant polypeptide is preferably the mature protein, or a
CC biologically active fragment, variant or derivative, that is capable
CC of eliciting an immune response, providing protection against one or
CC more strains of H. paragallinarum in chickens. Also claimed are
CC methods of using the haemagglutinin polypeptides and nucleic acids
CC for detection and diagnosis of infectious coryza in chickens.

XX Sequence 344 AA;

Query Match 9.0%; Score 86; DB 23; Length 344;
Best Local Similarity 20.1%; Pred. No. 0.81;
Matches 37; Conservative 31; Mismatches 72; Indels 44; Gaps 7;

QY 16 TLKLPAYWELSGFHLTDQWAIHYSY-----KYTEWSRKFELRGKYQDG 59
DB 55 TLRSVTYGVGGYQITDNFAVELGVDYDFGRAXXRGXGTVXXXTNHGAHLSLKASYPV 114
QY 60 SGYEAFK-----KEEYKDNSRFAIGTT--YSLNDALTLAGLAYDKAASKTHLSASIP 111
DB 115 EGLDVYARVCAALIRSDYKPTKRAAPNXTHELSKVPVFGAGLEYN-----LPSLP 166
QY 112 D-----TDRMWS-----IGATYKFTPNL--SVDVGFALRGKKKHVFETONIKGLLLVE 159
DB 167 ELALRVEYQWNVKVRXXXDGRVDYTPSIGSVTAGLSYRFGOSAPVPEKVVAKTFALN 226

QY 160 ADVT 163

DB 227 SDVT 230

RESULT 39

AAM92998

ID AAM92998 standard; protein; 340 AA.

XX AAM92998;

XX 19-MAY-1999 (first entry)

XX E. coli OmpF porln protein.

KW OmpF; porln; amphiphilic alpha-structure; beta-structure; GLUT;
KW membrane protein; glucose transporter protein; function.

XX Escherichia coll.

XX WO9618957-A1.

XX 20-JUN-1996.

XX 13-DEC-1995; 95WO-US16126.

XX 14-DEC-1994; 94US-0355844.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Cheung M, Czegledy F, Flschbarg J, Iserovich P;
PI LI J;

XX WPI; 1996-300839/30.

XX Predicting tendency to form amphiphilic alpha and beta structures -
PT using a novel algorithm to calculate values for subsequent graphical
PT analysis to predict protein structure

XX Disclosure; Fig 2; 106pp; English.

XX This invention describes a novel method for predicting the tendency of
CC a protein to form either: (a) an amphiphilic alpha-structure or (b) an
CC amphiphilic beta-structure. The methods are used particularly for
CC predicting the structure of membrane proteins such as glucose transporter
CC proteins (GLUTs). They can be used to discern the function of proteins.
CC They can also be used for the rational design or identification of
CC compounds which interact with the proteins or to engineer proteins having
CC particular structures. This sequence represents an Escherichia coli
CC OmpF porln which is used to illustrate the method of the invention.

XX Sequence 340 AA;

Query Match 8.9%; Score 85.5; DB 17; Length 340;

Best Local Similarity 25.2%; Pred. No. 0.9;

Matches 36; Conservative 22; Mismatches 60; Indels 25; Gaps 6;

QY 4 GVLGPYIGKSLTLKLPAYWELSGFHLTDQW-----IHYSKYTEWSRKFELRG 54

DB 186 GIVGAY-GAADRNLQEA--QPLGNGKKAQWATGLKYDANNIYLAANYGETRNPITN 242

QY 55 KYODGSGYEAFKKEEYKDNSRFAIGTYSLSNDALTLAGLAYDKAASKTHLSASIPDT 114

DB 243 KFTNTSGFANKTQDVLIVAOYQDFG-----LRPSIAYTKSKAKD--VEGIGD 290

QY 115 RM-WYSIGATYKFTPNLSVDVGF 136

DB 291 LVNFEVGATYYFNKNMSTYVDY 313

RESULT 40

AAB82612

ID AAB82612 standard; Protein; 341 AA.

XX AAB82612;

XX 02-OCT-2001 (first entry)

XX E. coli outer membrane protein ompF.

XX Outer membrane protein; ompF; structural protein; purification.

XX Escherichia coll.

XX WO200153333-A1.

XX 26-JUL-2001.

XX 01-NOV-2000; 2000WO-US30086.

XX 20-JAN-2000; 2000US-0490291.

XX (WELL/) MELLO C M.

XX (ARCI/) ARCIDIACONO S.

XX (BUTL/) BUTLER M M.

XX (USSA) US SEC OF ARMY.

XX Mello CM, Arcidiacono S, Butler MM;

XX WPI; 2001-483136/52.

XX Recovering structural polypeptides in a biological sample, useful for
PT purifying and spinning spider silks and other structural proteins,
PT comprises treating the sample containing the polypeptides with an acid

XX Claim 2; Page 43-44; 49pp; English.

XX The present sequence is that of the Escherichia coli ompF outer
CC membrane protein. The invention provides methods for purifying
CC structural proteins. Organic acids are used to lyse recombinant

CC cells or other biological samples (such as non-recombinantly
CC derived cells), and enrich the purity and yields of structural
CC proteins by hydrolysing many of the macromolecules while leaving
CC the structural proteins intact. In the present case, lyophilised
CC pellets of *E. coli* were lysed in 9.2 N valeric acid for 1 hr at
CC room temperature. The cell lysate was clarified by centrifugation
CC and applied to an SDS polyacrylamide gel for electrophoresis. The
CC ompF protein was then blotted onto a nitrocellulose membrane for
CC N-terminal sequencing. The simple extraction procedure yielded
CC ompF in approximately 85% purity. The new method has the
CC following advantages over prior art: it involves fewer steps,
CC requires less time and smaller volumes of reagents, results in
CC better recovery of protein at higher purity (70-99%), is easy to
CC scale up, and fibres can be spun in an environmentally benign
CC solution reducing hazardous waste accumulation and cost.

XX
SQ Sequence 341 AA;

Query Match 8.9%; Score 85.5; DB 22; Length 341;
Best Local Similarity 25.2%; Pred. NO. 0.9;
Matches 36; Conservative 22; Mismatches 60; Indels 25; Gaps 6;

Qy 4 GVLGPIYIGKSLTLKLPAYWELSGFHQLTDQWA-----IHYSKYKTEWSREKELRG 54
Db 186 GIVGAY-GAADRTN-QEA--OPLGNGKRAEQWATGLKYDANNIYLAANYGETRNATPTN 242
Qy 55 KYDGSGYEAFKKKEEYKDNSRFAIGTTYSLNDALTLRAGLAYDKAASKTHLSASIPDPTD 114
Db 243 KFTNTSGFANKTDQVLLVAQYQDFG-----LRPSIATYTSKAKD--VEGIGDVD 290
Qy 115 RM-WYSIGATYKFTPNLSVDVGF 136
Db 291 LVNYPEVGATYYFNKNMSTYVDY 313

Search completed: May 12, 2003, 09:40:18
Job time : 39 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: May 12, 2003, 09:41:20 ; Search time 20 seconds
(without alignments)
823.630 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 QHNGVLGPYIGKSLTKLP.....ADYTTKATANLYGLNLNRYF 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	133.5	13.9	456	10	US-09-997-664-92
2	83.5	8.7	261	10	US-09-765-272-152
3	81.5	8.5	452	10	US-09-815-242-5141
4	80	8.4	164	9	US-09-854-133-203
5	80	8.4	164	10	US-09-738-973-203
6	80	8.4	698	9	US-09-935-642-6
7	78	8.2	367	10	US-09-881-752A-188
8	76	7.9	878	9	US-10-254-534-2
9	74	7.7	282	12	US-10-062-994-6
10	74	7.7	282	12	US-10-062-994-6
11	73.5	7.7	1698	9	US-10-080-943-4
12	72.5	7.6	406	10	US-09-925-297-517
13	72	7.5	1434	9	US-10-080-505-9
14	70.5	7.4	375	9	US-09-984-130-38
15	70.5	7.4	383	9	US-09-905-291A-261
16	70.5	7.4	383	9	US-09-902-853-261
17	70.5	7.4	383	9	US-09-907-824-261
18	70.5	7.4	383	9	US-09-907-841-261
19	70.5	7.4	383	9	US-09-904-011-261

20	70.5	7.4	383	9	US-10-174-590-50	Sequence 50, Appl
21	70.5	7.4	383	9	US-10-176-758-50	Sequence 50, Appl
22	70.5	7.4	383	9	US-10-175-737-50	Sequence 50, Appl
23	70.5	7.4	383	9	US-09-906-742-261	Sequence 261, App
24	70.5	7.4	383	9	US-10-173-706-50	Sequence 50, Appl
25	70.5	7.4	383	9	US-10-175-738-50	Sequence 50, Appl
26	70.5	7.4	383	9	US-10-175-752-50	Sequence 50, Appl
27	70.5	7.4	383	9	US-10-176-482-50	Sequence 50, Appl
28	70.5	7.4	383	9	US-10-176-757-50	Sequence 50, Appl
29	70.5	7.4	383	9	US-10-176-913-50	Sequence 50, Appl
30	70.5	7.4	383	9	US-10-180-552-50	Sequence 50, Appl
31	70.5	7.4	383	9	US-10-180-557-50	Sequence 50, Appl
32	70.5	7.4	383	9	US-09-906-838-261	Sequence 261, App
33	70.5	7.4	383	9	US-09-907-613-261	Sequence 261, App
34	70.5	7.4	383	9	US-09-907-942-261	Sequence 261, App
35	70.5	7.4	383	9	US-10-173-700-50	Sequence 50, Appl
36	70.5	7.4	383	9	US-10-174-572-50	Sequence 50, Appl
37	70.5	7.4	383	9	US-10-174-579-50	Sequence 50, Appl
38	70.5	7.4	383	9	US-10-174-582-50	Sequence 50, Appl
39	70.5	7.4	383	9	US-10-174-588-50	Sequence 50, Appl
40	70.5	7.4	383	9	US-10-175-739-50	Sequence 50, Appl
41	70.5	7.4	383	9	US-10-175-740-50	Sequence 50, Appl
42	70.5	7.4	383	9	US-10-175-743-50	Sequence 50, Appl
43	70.5	7.4	383	9	US-10-176-488-50	Sequence 50, Appl
44	70.5	7.4	383	9	US-10-176-492-50	Sequence 50, Appl
45	70.5	7.4	383	9	US-10-176-747-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-09-997-664-92
; Sequence 92, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arie
; APPLICANT: Cattermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sarlasiani, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; FILE REFERENCE: BC1018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/585,174
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 92
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Pseudomonas mendocina KR-1
US-09-997-664-92

Query Match 13.9%; Score 133.5; DB 10; Length 456;

Best Local Similarity 26.7%; Pred. No. 1.9e-05;
Matches 40; Conservative 26; Mismatches 75; Indels 9; Gaps 4;

QY 13 GSILTK---LPAYWELSGFHLTDQWAIHYSVKYTWRSR-FKELRGKYQDGSYGAEFTKK 68

Db 291 GDIRVKDFEMPSALTFGFAHQFNERLWVAADYKRVVWSDVMEDISVDFKSSQSGGIDELP 350

QY 69 EBYKDNSRFAIGTYSNLDAITLRLAGLAYDKAASKTHLSAS-IPDTRMWSYIGATYKFT 127

Db 351 HNYQDITVASIGTAYRVNDKLTLAGYSVAOQALDSRLJLPVPAVLKHKHVSIGSDYSFD 410

QY 128 P----NLSVDVGFALHURGKKKHVFETQNIK 153

Db 411 KKSCLNLAIISFGLKESLNTSPSYLSGTETLK 440

RESULT 6
US-09-935-642-6
: Sequence 6, Application US/09935642
: Publication No. US20030044795A1
: GENERAL INFORMATION:
: APPLICANT: BYRJALSEN, Inger
: APPLICANT: LARSEN, Peter
: APPLICANT: STEPHEN, John
: TITLE OF INVENTION: Biochemical Markers for the Human
: TITLE OF INVENTION: Endometrium
: FILE REFERENCE: 8969-014
: CURRENT APPLICATION NUMBER: US/09/935,642

Query Match	8.2%;	Score 78;	DB 10;	Length 367;
Best Local Similarity	26.5%;	Pred. No. 5.9;		
Matches 35;	Conservative 21;	Mismatches 46;	Indels 30;	Gaps 7;
Qy	63	EAPTKKEEYKONSR-FATGTTTYSLNDALTLRAGLAYDKAASKTHLSAS1PDTDRMYS--	119	
Db	136	EPITNPLEFEENLRLEVQFAQSQRML---SLSQIAAISNSLNALDPNS-----YSKN	188	
Qy	120	IGATYKFTPNLSVDVGFQFAHLRKKKHFEVQNIKGILLLVAEADYTT-----KAT	167	
Db	189	ISSMY-----GVSLSVGYKHFPTKKN---QGLRYLYFYDYGYTTFGFGVNGFDGLGKMN	240	
Qy	168	ANLYGLNLNYRF	179	
Db	241	NHLYGLGIDYLY	252	

```

; SEQ ID NO 6
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Anaplasma marginale
US-10-062-994-6

Query Match
Best Local Similarity 7.7%; Score 74; DB 12; Length 282;
Matches 49; Conservative 27; Mismatches 86; Indels 38; Gaps 11;

QY 14 SLTLKLPAYWELSGF-----HQLTD-QWAIHYSY-----KYTEWSRFEKELRGKYQDGS 60
DB 84 TIDVSPANFSGKYTFASFKNLITSDGAVGYSLGARVELEASRYRFRATLADGOYAKS 143
QY 61 GYE---AFTKKEEYKDNRSRFAIGTTYSLNDALTLAGLAYDKAASKTHLS-----AS 109
DB 144 GAESLAATRDANITETNYFVVKIDEITNTSVNLNG--CYDVLHTDLPVSPYCAIGIGAS 201
QY 110 IPD-----TDRMWY--SIGATYKFTPNLS--VDVGFH--LRGKKKHVFETQNIKGLLVE 159
DB 202 FVDISKQVTTKLAYRGKVGISYQFTPEISLVAGGFVHGLFDESFKDIPAHNSVK--FSGE 259
QY 160 ADYTTKATANLYGLNLNRYF 179
DB 260 AKASVKAHIADYGFNLGARF 279

RESULT 10
US-10-062-994-6
; Sequence 6, Application US/10062994
; Patent No. US20020132789A1
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Bowle, Michael V.
; APPLICANT: Burrige, Michael J.
; APPLICANT: Mahan, Suman M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Kurangirwa, Fred R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Whitmire, William M.
; APPLICANT: Allemen, Arthur R.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases and
; FILE REFERENCE: UF-167XC3
; CURRENT APPLICATION NUMBER: US/10/062.994
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/533,662
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/337,827
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/953,326
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 08/733,230
; PRIOR FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Anaplasma marginale
US-10-062-994-6

Query Match
Best Local Similarity 7.7%; Score 74; DB 12; Length 282;
Matches 49; Conservative 27; Mismatches 86; Indels 38; Gaps 11;

QY 14 SLTLKLPAYWELSGF-----HQLTD-QWAIHYSY-----KYTEWSRFEKELRGKYQDGS 60
DB 84 TIDVSPANFSGKYTFASFKNLITSDGAVGYSLGARVELEASRYRFRATLADGOYAKS 143
QY 61 GYE---AFTKKEEYKDNRSRFAIGTTYSLNDALTLAGLAYDKAASKTHLS-----AS 109
DB 144 GAESLAATRDANITETNYFVVKIDEITNTSVNLNG--CYDVLHTDLPVSPYCAIGIGAS 201
QY 110 IPD-----TDRMWY--SIGATYKFTPNLS--VDVGFH--LRGKKKHVFETQNIKGLLVE 159
DB 202 FVDISKQVTTKLAYRGKVGISYQFTPEISLVAGGFVHGLFDESFKDIPAHNSVK--FSGE 259
QY 160 ADYTTKATANLYGLNLNRYF 179
DB 260 AKASVKAHIADYGFNLGARF 279

; SEQ ID NO 2
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:beII gene (branching enzyme II)
US-10-254-534-2

Query Match
Best Local Similarity 7.9%; Score 76; DB 9; Length 878;
Matches 16; Conservative 10; Mismatches 18; Indels 2; Gaps 1;

QY 38 HVSRYKTEWSRFEKELRGKYQDGSYEATKKEEYKDNRSRFAIGTTY 83
DB 193 HLDYRSQYKUREAIDYE--GGLEAFSRGYEKMGMFTRSATGITY 236

RESULT 9
US-10-062-994-6
; Sequence 6, Application US/10062994
; Patent No. US2001001661A1
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Bowle, Michael V.
; APPLICANT: Burrige, Michael J.
; APPLICANT: Mahan, Suman M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Kurangirwa, Fred R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Whitmire, William M.
; APPLICANT: Allemen, Arthur R.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases and
; FILE REFERENCE: UF-167XC3
; CURRENT APPLICATION NUMBER: US/10/062.994
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/533,662
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/337,827
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/953,326
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 08/733,230
; PRIOR FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
```

Db 144 GAESLAAITRDANITETNYFVVKIDETITNSVNLNG--CYDVLHLDLPVSPYVCAGIGAS 201
QY 110 IPD-----TDRMY--SIGATYKFTNLS--VDVGFH--LRGKKKHFVETQNIKGLLLVE 159
Db 202 FVDISKQVTKLAYRGKVGISQFTPEISLVAGCFYHGLFDESYKDIPAHNSVK--FSGE 259
QY 160 ADYTTKATANLYGLNLNRYF 179
Db 260 AKASVRAHIADYGFNLGARE 279
RESULT 11
US-10-080-943-4
; Sequence 4, Application US/10080943
; Publication No. US20030073236A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Tsai, Shih-Chong
; TITLE OF INVENTION: P193 PROTEINS AND NUCLEIC ACIDS, AND USES THEREOF
; FILE REFERENCE: IU99-PCT
; CURRENT APPLICATION NUMBER: US/10/080.943
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/150.266
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-943-4
Query Match 7.7%; Score 73.5; DB 9; Length 1698;
Best Local Similarity 23.9%; Pred. No. 1.2e+02;
Matches 28; Conservative 17; Mismatches 47; Indels 25; Gaps 3;
QY 64 APTKKEEYKDNRSFATGTTYSNLDTLRLAGLAYDKAAKTHLSA-----SIPD 112
Db 350 SFRSRFRPRSEFASGNTYALYVRTLQPGMRVRLDDYEISAGDEGEFGCSNGVPP 409
QY 113 TDRMYSIGATYKFTNLSVDVGFHRLGKKKHFVETQNIKGLLLVEADYTTKATN 169
Db 410 VOVFWESTGRYVWHWMLILGF-----EEDIED--MVEADEYQGAVAS 452
RESULT 12
US-09-925-297-517
; Sequence 517, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925.297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124.270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 517
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-517
Query Match 7.6%; Score 72.5; DB 10; Length 406;
Best Local Similarity 21.7%; Pred. No. 24;
Matches 45; Conservative 25; Mismatches 66; Indels 71; Gaps 12;
QY 18 KLPAYWELSGFHQLDQMAHYKYYTMSRFRKELRGKYODGSGYEAFYTK--EYKDNS 75

Db 173 KRPAIWIDTGIH--SREWVTOAS---GVW-----FAKKITQDYQDA 209
QY 76 REA--TGTYSNLDTLRLAGLAYDKAA-----SKTHLSASI---PDTDRMYS----- 119
Db 210 AFTAILDTLDIFLEIVTNDGFAFTHTNRMRKTKTSHTAGSLCIGVDPRNRWDAGFGLS 269
QY 120 -----IGATY--KFTPNLSVD-----VGFHRLGKKKHFVETQNIKGLLLVEADYTTK-- 165
Db 270 GASSNPCSEYTHGKFA-NSEVEVKISVDFVKDHGNIKAFISHSYQLLMYPYGYKTEPV 328
QY 166 -----ATANLYGLNLNRY 177
Db 329 PQDELDQLSKAAVATLASLYGTFKNY 355
RESULT 13
US-10-080-505-9
; Sequence 9, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCE/DHR
; CURRENT APPLICATION NUMBER: US/10/080.505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296.791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839.996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (568)..(568)
; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
US-10-080-505-9
Query Match 7.5%; Score 72; DB 9; Length 1434;
Best Local Similarity 19.1%; Pred. No. 1.3e+02;
Matches 34; Conservative 30; Mismatches 56; Indels 58; Gaps 9;
QY 46 WSRFKELRGKYODGSGYEAFYTKKEEYK-----DNSRFAIGTTYSLND----- 87
Db 1189 WTIADQKRRY-DSDAFRAYQOKTNLRQIGVQKALDNGRICAVFSHSRSDNTEDEQVKNH 1247
QY 88 -ALTLRAGLAY-----DKAASKTHLSASIPDTRMYSICATY 124
Db 1248 ATLTMSGFGAYQYQWGDLOFGVNVGAGISASKMAEESRKTHRKA-----INGVNWSY 1300
QY 125 KFTPNLSVDVGFHRLGKKKHFVETQNIKGLLLVEADYTTKATA-NLY--GLNLNRYF 179
Db 1301 QFRIG---QLGIQPYLGVNRYFIERENYQS---EEVKVQTPSLAFNRYNAGIRVDYTF 1352
RESULT 14
US-09-984-130-38
; Sequence 38, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984.130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243.792
; PRIOR FILING DATE: 2000-10-30

;; PRIOR APPLICATION NUMBER: 09/836,353
;; PRIOR FILING DATE: 2001-04-18
;; PRIOR APPLICATION NUMBER: 60/198,407
;; PRIOR FILING DATE: 2000-04-19
;; PRIOR APPLICATION NUMBER: PCT/US99/25031
;; PRIOR FILING DATE: 1999-10-27
;; PRIOR APPLICATION NUMBER: 60/105,971
;; PRIOR FILING DATE: 1998-10-28
;; NUMBER OF SEQ ID NOS: 149
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 38
;; LENGTH: 375
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-984-130-38

Query Match 7.4%; Score 70.5; DB 9; Length 375;
Best Local Similarity 23.8%; Pred. No. 35;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
QY 7 GPYIGKSLTLKLPAYWELSGFHLTDQWATHYSKYKTEMSRFEKELRGKY----- 56
Db 66 GPQCHKGT---PLTYEAKQY-----LSYETLVANGSRTETQVGIYILSSGDDGAQ 114
QY 57 --QDGSYEAFTEKKEKYDMSRFAI-CTTYSLN-----DALTLRAGLAYDKAAKTHLSAS 109
Db 115 HRDSSGSGKSRKRRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGTGCTGLVAEKHVLTA 174
QY 110 --IPDTRMYSICATY-KETPNLSV-----DVGFAHLRGKK 143
Db 175 HCHTD-----GKTVYKGTQLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVETQNIKGLLVEADYTTKATANLYGLNLN 177
Db 227 THVP-----KGWI-----KGNANDIGMDYD 247

RESULT 15
US-09-905-291A-261
;; Sequence 261, Application US/09905291A
;; Patent No. US20020160374A1
;; GENERAL INFORMATION:
;; APPLICANT: Genentech, Inc.
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Flivaroff, Ellen
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, A.
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, Christopher J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth, J.
;; APPLICANT: Kijavini, Ivar J.
;; APPLICANT: Mather, Jennie P.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William, I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: 10466-14
;; CURRENT APPLICATION NUMBER: US/09/905,291A
;; CURRENT FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: PCT/US00/04414

;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: US 60/143,048
;; PRIOR FILING DATE: 1999-07-07
;; PRIOR APPLICATION NUMBER: US 60/145,698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: US 60/146,222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/20944
;; PRIOR FILING DATE: 1999-09-13
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 261
;; LENGTH: 383
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-905-291A-261

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
QY 7 GPYIGKSLTLKLPAYWELSGFHLTDQWATHYSKYKTEMSRFEKELRGKY----- 56
Db 66 GPQCHKGT---PLTYEAKQY-----LSYETLVANGSRTETQVGIYILSSGDDGAQ 114
QY 57 --QDGSYEAFTEKKEKYDMSRFAI-CTTYSLN-----DALTLRAGLAYDKAAKTHLSAS 109
Db 115 HRDSSGSGKSRKRRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGTGCTGLVAEKHVLTA 174
QY 110 --IPDTRMYSICATY-KETPNLSV-----DVGFAHLRGKK 143
Db 175 HCHTD-----GKTVYKGTQLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVETQNIKGLLVEADYTTKATANLYGLNLN 177
Db 227 THVP-----KGWI-----KGNANDIGMDYD 247

RESULT 16
US-09-902-853-261
;; Sequence 261, Application US/09902853
;; Publication No. US20020192659A1
;; GENERAL INFORMATION:
;; APPLICANT: Genentech, Inc.
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Flivaroff, Ellen

; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 261
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-261

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
Qy 7 GPYIGKSLTLKLPAYWELSGFHLTDQWATHYSYKYTEWSRFEKELRGKY-----LSYETLYANGSRTEQVGIYILSSSGDGAQ 114
Db 66 GPQCHKGT---PLTYEAKQY-----LSYETLYANGSRTEQVGIYILSSSGDGAQ 114
Qy 57 --QDGSYEAPTKKEEYKDNRSFPAI-GTYSLSN-----DALTLRAGLAYDKAAKTHLSAS 109
Db 115 HRDSSGSKSRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGTCTTLVAERHVLTA 174
Qy 110 --IPDTRMYSICATY-KETPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTVKGTQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
Qy 144 KHFEVTONIKGLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 18

US-09-907-841-261
; Sequence 261, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 261
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-841-261

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
Qy 7 GPYIGKSLTLKLPAYWELSGFHLTDQWATHYSYKYTEWSRFEKELRGKY-----LSYETLYANGSRTEQVGIYILSSSGDGAQ 114
Db 66 GPQCHKGT---PLTYEAKQY-----LSYETLYANGSRTEQVGIYILSSSGDGAQ 114
Qy 57 --QDGSYEAPTKKEEYKDNRSFPAI-GTYSLSN-----DALTLRAGLAYDKAAKTHLSAS 109
Db 115 HRDSSGSKSRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGTCTTLVAERHVLTA 174
Qy 110 --IPDTRMYSICATY-KETPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTVKGTQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
Qy 144 KHFEVTONIKGLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 19

US-09-904-011-261
; Sequence 261, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/904,011
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 261
LENGTH: 383
TYPE: PRT
ORGANISM: Homo Sapien
US-09-904-011-261

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSYKYTEWSRFEKLRGKY-----56
Db 66 GPQCHKGT---PLPTYEAKQY-----LSYETLYANGSRRTETQVGIYILSSSGDGAQ 114
QY 57 --QDGSYGAEFTKKEEYKONSRAFI-GTTYSLN-----DALTLRAGLAYDKAAASKTHLSAS 109
Db 115 HRDSSGSGKRRKRIQYGYDSRFSIFGKDFLLNYPSTSVKLTGCTGTGLVAEKHVLTA 174
QY 110 --IPDTRMWSYGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIIHD-----GKTYVGKTKLRVGLPKPKDGRGNDSTSAMPEOMKFWLRVKR 226
QY 144 KHIVETQNIKGLLLVEADYTTATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247
US-09-904-011-261

RESULT 20
US-10-174-590-50
Sequence 50, Application US/10174590
Publication No. US20030008352A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C42
CURRENT APPLICATION NUMBER: US/10/174,590
CURRENT FILING DATE: 2002-06-18
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 50
LENGTH: 383
TYPE: PRT
ORGANISM: Homo Sapien
US-10-174-590-50
Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;
QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSYKYTEWSRFEKLRGKY-----56
Db 66 GPQCHKGT---PLPTYEAKQY-----LSYETLYANGSRRTETQVGIYILSSSGDGAQ 114
QY 57 --QDGSYGAEFTKKEEYKONSRAFI-GTTYSLN-----DALTLRAGLAYDKAAASKTHLSAS 109
Db 115 HRDSSGSGKRRKRIQYGYDSRFSIFGKDFLLNYPSTSVKLTGCTGTGLVAEKHVLTA 174
QY 110 --IPDTRMWSYGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIIHD-----GKTYVGKTKLRVGLPKPKDGRGNDSTSAMPEOMKFWLRVKR 226
QY 144 KHIVETQNIKGLLLVEADYTTATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247
RESULT 21
US-10-176-758-50
Sequence 50, Application US/10176758
Publication No. US20030008353A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
CURRENT FILING DATE: 2002-06-21

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-50

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYTGKSLTLKLPAYWELSGFHLTDQWATHYSKYKTEWSRFRKELRGKY-----56
Db 66 GPQCHKGT---PLPTYEAKQY-----LSYETLYANGSRRTETQVGIIYLSSSGGGAQ 114
QY 57 --QDGSYEATTKKEEYKDNRSFAL-GTYSLSN-----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKSRKKQIYGYDSRFSIEGKDFLLNYPFSTSVKLSGCTGLVAEKHVLTA 174
QY 110 --IPDTRMWSIGATY-KFTPNLSV-----DVGFALRGKK 143
Db 175 HCIHD-----GKTYVKGTQKLRVGLPKPKFKDGGGRANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVETQNIKGLLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 22

US-10-175-737-50
; Sequence 50, Application US/10175737
; Publication No. US20030013153A1

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-50

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYTGKSLTLKLPAYWELSGFHLTDQWATHYSKYKTEWSRFRKELRGKY-----56
Db 66 GPQCHKGT---PLPTYEAKQY-----LSYETLYANGSRRTETQVGIIYLSSSGGGAQ 114
QY 57 --QDGSYEATTKKEEYKDNRSFAL-GTYSLSN-----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKSRKKQIYGYDSRFSIEGKDFLLNYPFSTSVKLSGCTGLVAEKHVLTA 174
QY 110 --IPDTRMWSIGATY-KFTPNLSV-----DVGFALRGKK 143
Db 175 HCIHD-----GKTYVKGTQKLRVGLPKPKFKDGGGRANDSTSAMPEQMKFQWIRVKR 226

QY 144 KHFEVETQNIKGLLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 23

US-09-906-742-261
; Sequence 261, Application US/09906742
; Publication No. US20030023054A1

; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Oiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

```

QY 110 --IPDTRDMWYSTGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HC1HD-----GKTYVGKQTKLRVGLFKPKFKDGGRGANDSTSAMPEOMKQFQWIRVKR 226
QY 144 KHEVETQNIKGLLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYD 247

RESULT 25
US-10-175-738-50
; Sequence 50, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P34301C45
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-50

Query Match 7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps

QY 7 GPYIGKSLTKLPAYWELSGFHQLTDQNAIHYSYKYTENSFKELRGKY----- 56
Db 66 GPQCHRGKT--PLPTVEAAKQY-----LSVETLYANGSRPTQVGIYILSSSGDGAQ 114
QY 57 --QDGSYGAEFTKKEEKONSREAI-GTYSLSN----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSGSKRRKQRYGVDSRESIFGKDFLLNYPFSTSVKLSGTGCTGLVAEKHVLTA 174
QY 110 --IPDTRDMWYSTGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HC1HD-----GKTYVGKQTKLRVGLFKPKFKDGGRGANDSTSAMPEOMKQFQWIRVKR 226
QY 144 KHEVETQNIKGLLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYD 247

RESULT 26
US-10-175-752-50
; Sequence 50, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.

```



```

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,838
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 261
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-261
Query Match          7.4%   Score 70.5;   DB 9;   Length 383;
```

```

Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSCFHQLTDQWAIHYSKYKTEWSRPFELRGKY----- 56
Db 66 GPQCHKGT---PLPTYBEAKQY-----LSYETLYANGSRRTETQVGIVILSSSGDGAQ 114
QY 57 --QDGSYEAFTKKEEYKDNRSFAI-GTTYSLN----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSGKSRKQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGTGCTGLVAEKHVLTA 174
QY 110 --IPDTRMYSIGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTVKGTQKLRVGLPKFKDGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVETQNIKLLLEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 33
US-09-907-613-261
; Sequence 261, Application US/09907613
; Publication No. US20030027145A1.
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
```


; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 261
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-907-613-261

Query Match
Best Local Similarity 7.4%; Score 70.5; DB 9; Length 383;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSYKYTEMSRKELRGKY-----56
Db 66 GPOCHKGT---PLPTVEEAKQY-----LSYETLYANGSRVETQVGIYILSSGGGGAQ 114

QY 57 --ODSGGYEAFKKEEYKDNRSFAI-GTYYSLN-----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKRRKRQIYGYDSRFSIFGKDFLLNYPSTSVKLSGTGCTGLVAEKHVLTA 174

QY 110 --IPDTRMWSYGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCHD-----GKTVVKGTKLRVGLFKPKFKDGGRGANDSTSAMPEQMKFQWIRVKR 226

QY 144 KHEVETQNIKGLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 34
US-09-907-942-261
; Sequence 261, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivár J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,942
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 261
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-907-942-261

Query Match
Best Local Similarity 7.4%; Score 70.5; DB 9; Length 383;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSYKYTEMSRKELRGKY-----56
Db 66 GPOCHKGT---PLPTVEEAKQY-----LSYETLYANGSRVETQVGIYILSSGGGGAQ 114

QY 57 --ODSGGYEAFKKEEYKDNRSFAI-GTYYSLN-----DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKRRKRQIYGYDSRFSIFGKDFLLNYPSTSVKLSGTGCTGLVAEKHVLTA 174

QY 110 --IPDTRMWSYGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCHD-----GKTVVKGTKLRVGLFKPKFKDGGRGANDSTSAMPEQMKFQWIRVKR 226

QY 144 KHEVETQNIKGLLVEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 35
US-10-173-700-50
; Sequence 50, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-50

Query Match
Best Local Similarity 7.4%; Score 70.5; DB 9; Length 383;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSRFKELRGKY----- 56
Db 66 GPQCHKGT---PLPTYEEAKQY-----LSYETLYANGSRTEQTQVGIIYILSSSGDGAQ 114
QY 57 --QDGSGYEAFTKKEEYKDNSRFAT-GTYSLSN---DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKSRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSCTGTGLVAEKHVLTA 174
QY 110 --IPDTRMWSIGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTYVKGTKLRVGLPKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVTONIKGLLLEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 36
US-10-174-572-50
; Sequence 50, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-572-50

Query Match
Best Local Similarity 7.4%; Score 70.5; DB 9; Length 383;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSRFKELRGKY----- 56
Db 66 GPQCHKGT---PLPTYEEAKQY-----LSYETLYANGSRTEQTQVGIIYILSSSGDGAQ 114
QY 57 --QDGSGYEAFTKKEEYKDNSRFAT-GTYSLSN---DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKSRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSCTGTGLVAEKHVLTA 174
QY 110 --IPDTRMWSIGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTYVKGTKLRVGLPKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVTONIKGLLLEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 37
US-10-174-579-50
; Sequence 50, Application US/10174579
; Publication No. US20030027264A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-579-50

Query Match
Best Local Similarity 7.4%; Score 70.5; DB 9; Length 383;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSRFKELRGKY----- 56
Db 66 GPQCHKGT---PLPTYEEAKQY-----LSYETLYANGSRTEQTQVGIIYILSSSGDGAQ 114
QY 57 --QDGSGYEAFTKKEEYKDNSRFAT-GTYSLSN---DALTLRAGLAYDKAASKTHLSAS 109
Db 115 HRDSSGSKSRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSCTGTGLVAEKHVLTA 174
QY 110 --IPDTRMWSIGATY-KFTPNLSV-----DVGFAHLRGKK 143
Db 175 HCIHD-----GKTYVKGTKLRVGLPKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHFEVTONIKGLLLEADYTTKATANLYGLNLNY 177
Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 38
US-10-174-572-50
; Sequence 50, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-572-50
```

```
US-10-174-582-50
; Sequence 50, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-582-50

Query Match      7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHOLTQDWAHYSKYKTEWSRPFKELRGKY-----56
DB 66 GPOCHKGT---PLPTYEAKQY-----LSYETLVANGSRRTETQVGIYILSSSGDGAQ 114
QY 57 --QDGSGEAFTKKEEYKDNRSRFAI-GTITYSLN----DALTLRAGLAYDKAAKSTHLSAS 109
DB 115 HRDSSGSGSKRRKQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGCTCTTLVAEKHVLTA 174
QY 110 --IPDTRMYSIGATY-KETPNLSV-----DVGFAHLRGKK 143
DB 175 HCIHD-----GKTVYKGTQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHVETONIKGLLVEADYTTKATANLYGLNLNY 177
DB 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 39
US-10-174-588-50
; Sequence 50, Application US/10174588
; Publication No. US20030027266A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C38
; CURRENT APPLICATION NUMBER: US/10/174,588
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
```

```
US-10-174-588-50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-588-50

Query Match      7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHOLTQDWAHYSKYKTEWSRPFKELRGKY-----56
DB 66 GPOCHKGT---PLPTYEAKQY-----LSYETLVANGSRRTETQVGIYILSSSGDGAQ 114
QY 57 --QDGSGEAFTKKEEYKDNRSRFAI-GTITYSLN----DALTLRAGLAYDKAAKSTHLSAS 109
DB 115 HRDSSGSGSKRRKQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGCTCTTLVAEKHVLTA 174
QY 110 --IPDTRMYSIGATY-KETPNLSV-----DVGFAHLRGKK 143
DB 175 HCIHD-----GKTVYKGTQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHVETONIKGLLVEADYTTKATANLYGLNLNY 177
DB 227 THVP-----KGWI-----KGNANDIGMDYDY 247

RESULT 40
US-10-175-739-50
; Sequence 50, Application US/10175739
; Publication No. US20030027267A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C46
; CURRENT APPLICATION NUMBER: US/10/175,739
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 50
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-739-50

Query Match      7.4%; Score 70.5; DB 9; Length 383;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 51; Conservative 22; Mismatches 66; Indels 75; Gaps 11;

QY 7 GPYIGKSLTLKLPAYWELSGFHOLTQDWAHYSKYKTEWSRPFKELRGKY-----56
DB 66 GPOCHKGT---PLPTYEAKQY-----LSYETLVANGSRRTETQVGIYILSSSGDGAQ 114
QY 57 --QDGSGEAFTKKEEYKDNRSRFAI-GTITYSLN----DALTLRAGLAYDKAAKSTHLSAS 109
DB 115 HRDSSGSGSKRRKQIYGYDSRFSIFGKDFLLNYPFSTSVKLSGCTCTTLVAEKHVLTA 174
QY 110 --IPDTRMYSIGATY-KETPNLSV-----DVGFAHLRGKK 143
DB 175 HCIHD-----GKTVYKGTQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKR 226
QY 144 KHVETONIKGLLVEADYTTKATANLYGLNLNY 177
DB 227 THVP-----KGWI-----KGNANDIGMDYDY 247
```

Db 227 THVP-----KGWI-----KGNANDIGMDYDY 247

Search completed: May 12, 2003, 09:46:36
Job time : 25 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:39:14 ; Search time 17 Seconds
(without alignments)
1012.239 Million cell updates/sec

Title: US-09-995-493-52
Perfect score: 957
Sequence: 1 QHNGVLGPYIGKSLTKLP.....ADYTTKATANLYGLNLNRYF 179
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	570.5	59.6	450	2 B40183	outer membrane pro
2	559	58.4	455	2 A40183	outer membrane pro
3	549	57.4	459	2 A28787	outer membrane pro
4	549	57.4	459	2 F84065	outer membrane pro
5	421	44.0	412	2 G82406	long-chain fatty a
6	372.5	38.9	423	2 A40334	probable long-chain
7	338	35.3	437	2 A80805	long-chain fatty a
8	334	34.9	432	2 B82250	long-chain fatty a
9	332	34.7	448	2 F65007	long-chain fatty a
10	332	34.7	448	2 C81032	long-chain fatty a
11	332	34.7	448	2 D85876	long-chain fatty a
12	267.5	28.0	428	2 B82248	long-chain fatty a
13	207	21.6	447	2 E82729	outer membrane pro
14	194.5	20.3	432	2 E87508	outer membrane pro
15	190.5	19.9	463	2 A83073	probable outer mem
16	167	17.5	424	2 H83484	probable outer mem
17	144.5	15.1	466	2 A81239	outer membrane pro
18	143.5	15.0	532	2 G83424	hypothetical prote
19	142.5	14.9	464	2 H82011	probable outer mem
20	132.5	13.8	403	2 C70385	hypothetical prote
21	111	11.6	308	2 H70355	hypothetical prote
22	108	11.3	587	2 G84624	outer membrane pro
23	106.5	11.1	453	2 S53995	Tox protein - Pse
24	102	10.7	587	2 C71889	probable outer mem
25	100	10.4	753	2 D83081	probable outer mem
26	96.5	10.1	708	2 B81038	TonB-dependent rec
27	94	9.8	572	2 A81209	N-acetylmuramoyl-L
28	92.5	9.7	2425	2 D69426	surface layer prot
29	91.5	9.6	730	2 B83592	hypothetical prote

ALIGNMENTS

RESULT 1

B40183

outer membrane protein P1 precursor, subtype 6U - Haemophilus influenzae (strain 8358)
C:Species: Haemophilus influenzae
C:Date: 28-Oct-1992 #sequence_revision 30-Jan-1993 #text_change 12-Dec-1997
C:Accession: B40183
R:Musson Jr., R.; Grass, S.; Einhorn, M.; Bailey, C.; Newell, C.
Infect. Immun. 57, 3300-3305, 1989.
A:Title: Comparative analysis of the structures of the outer membrane protein P1 gene
A:Reference number: A40183; MUID:90035394; PMID:2572549
A:Accession: B40183
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <MUN>
A:CROSS-references: GB:M27683
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-450/Product: outer membrane protein P1 #status predicted <MAT>

Query Match 59.6%; Score 570.5; DB 2; Length 450;
Best Local Similarity 62.7%; Pred. No. 4.8e-45;
Matches 106; Conservative 30; Mismatches 30; Indels 3; Gaps 2;
QY 12 KGSLLIKLPAYWELSGFHQITDOWATHYKYYKTENSRFKELRGKYQDGSYEAFTKKEY 71
DB 284 KGNLTLLKLPDYLELSGFHQLTKFAVHYKYYKTHWSRLTKLHASFEDGK--KAFDKELQY 341
QY 72 KONSRAFTGTTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYSGATYKFTPNLS 131
DB 342 SNNSRIALGASYNLDEKLTLAGIAYDQAASRHRSAAIPDTRNRTWYSLGATYKFTPNLS 401
QY 132 VDVGFAHLRGKKHFVETQNIK-GLLLVEADYTTKATANLYGLNLNRYF 179
DB 402 VDLGVAYLKGKVKHFEQAQAGGFTTTTANTVTSQAHANLYGLNLNYSF 450

RESULT 2

A40183

outer membrane protein P1 precursor, subtype 3L - Haemophilus influenzae (strain 1613)
C:Species: Haemophilus influenzae
C:Date: 28-Oct-1992 #sequence_revision 30-Jan-1993 #text_change 20-Aug-1999
C:Accession: A40183
R:Musson Jr., R.; Grass, S.; Einhorn, M.; Bailey, C.; Newell, C.
Infect. Immun. 57, 3300-3305, 1989
A:Title: Comparative analysis of the structures of the outer membrane protein P1 gene
A:Reference number: A40183; MUID:90035394; PMID:2572549
A:Accession: A40183
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <MUN>
A:CROSS-references: GB:M63151; GB:M27682; NID:g148954; PIDN:AAA24991.1; PID:g148955

outer membrane pro
outer membrane pro
outer membrane uah
peptide methionine
hypothetical prote
probable PagC-like
outer membrane por
vitamin B12 recept
penicillin-binding
hypothetical prote
adhesion and penet
hemeoglobin-haptog
choleline binding pr
ADP-ribosylarginin
hypothetical prote

C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-455/Product: outer membrane protein P1 #status predicted <NAT>

Query Match 58.4%; Score 559; DB 2; Length 455;
Best Local Similarity 61.6%; Pred. No. 5.6e-44;
Matches 106; Conservative 29; Mismatches 31; Indels 6; Gaps 2;

QY 12 KGSITLKPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKELRGKYQDGSYGAEFTKKEEY 71
DB 286 KGNLTFTLPDYLELSGFHQLTDKFAVHYSKYKTHWSRLTKLHASFEDGK--KAPDKELQY 343
QY 72 KDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYSGATYKFTPNLS 131
DB 344 SNNSRVALGASYNLDEKLTLAGIAYDQAAASRHRSAAIPDTRTWYSLGATYKFTPNLS 403
QY 132 VDVGFAHLRGKKHVFETONIKG-----LLLEADYTTKATANLYGLNLNRYF 179
DB 404 VDLGAYLKGKKVHFKVETGDKRTLTNTANTYTSQAHANLYGLNLNYSF 455

RESULT 3
A28787

outer membrane protein P1 precursor - Haemophilus influenzae (type b)
C:Species: Haemophilus influenzae
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Aug-1999
C:Accession: A30510; A28787
R:Munson Jr., R.; Grass, S.
Infect. Immun. 56, 2235-2242, 1988
A:Title: Purification, cloning, and sequence of outer membrane protein P1 of Haemophilus
A:Reference number: A30510; MUID:88314258; PMID:2842261
A:Accession: A30510
A:Molecule type: DNA

A:Residues: 1-459 <MUN>
A:Cross-references: GB:J03381; NID:g148952; PIDN:AAA24990.1; PID:g148953
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-459/Product: outer membrane protein P1 #status predicted <NAT>

Query Match 57.4%; Score 549; DB 2; Length 459;
Best Local Similarity 61.0%; Pred. No. 4.8e-43;
Matches 105; Conservative 27; Mismatches 34; Indels 6; Gaps 2;

QY 12 KGSITLKPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKELRGKYQDGSYGAEFTKKEEY 71
DB 290 KGNLTFTLPDYLELSGFHQLTDKLAHYSKYKTHWSRLTKLHASFEDGK--KAPDKELQY 347
QY 72 KDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYSGATYKFTPNLS 131
DB 348 SNNSRVALGASYNLDEKLTLAGIAYDQAAASRHRSAAIPDTRTWYSLGATYKFTPNLS 407

QY 132 VDVGFAHLRGKKHVFETONIKG-----LLLEADYTTKATANLYGLNLNRYF 179
DB 408 VDLGAYLKGKKVHFKVETGDKRTLTNTANTYTSQAHANLYGLNLNYSF 459

RESULT 4
F64065

outer membrane protein P1 precursor - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Aug-1999
C:Accession: F64065
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64065
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-459 <TIGR>
A:Cross-references: GB:U32723; GB:L42023; NID:g1573363; PIDN:AAC22060.1; PID:g1573372
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-459/Product: outer membrane protein P1 #status predicted <NAT>

Query Match 57.4%; Score 549; DB 2; Length 459;
Best Local Similarity 61.6%; Pred. No. 4.8e-43;
Matches 106; Conservative 28; Mismatches 32; Indels 6; Gaps 3;

QY 12 KGSITLKPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKELRGKYQDGSYGAEFTKKEEY 71
DB 290 KGNLTFTLPDYLELSGFHQLTDKLAHYSKYKTHWSRLTKLNASFEDGK--KAPDKELQY 347
QY 72 KDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYSGATYKFTPNLS 131
DB 348 SNNSRVALGASYNLDEKLTLAGIAYDQAAASRHRSAAIPDTRTWYSLGATYKFTPNLS 407
QY 132 VDVGFAHLRGKKHVFETONIKG-----KGLLL-VEADYTTKATANLYGLNLNRYF 179
DB 408 VDLGAYLKGKKVHFKVETGDKRTLTNTANTYTSQAHANLYGLNLNYSF 459

RESULT 5
G82406

long-chain fatty acid transport protein VCA0862 [imported] - Vibrio cholerae (strain
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82406
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82406
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-412 <HEI>
A:Cross-references: GB:AE004414; GB:AE003853; NID:g9658293; PIDN:AAF96760.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

A:Gene: VCA0862
A:Map position: 2
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 44.0%; Score 421; DB 2; Length 412;
Best Local Similarity 46.9%; Pred. No. 3e-31;
Matches 83; Conservative 27; Mismatches 61; Indels 6; Gaps 2;

QY 3 NGVLGPVIGKSLTLKLPAYWELSGFHQLTDQWAIHYSKYKTEWSEFKELRGKYQDGSYG 62
DB 241 SGATSAVVPQGLKIELPAIWELSGFHQLNEQWAVHYSYQOTDSSFEELTATSSQCKNG 300
QY 63 EAFTKKEEYKDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYSGA 122
DB 301 TCFQKIEQYQDNGRWSVGATYTLNTTTLRAGLAFDEQAGKATL--SIPDSRDEWYSAGL 358

QY 123 TYKFTPNLSVDVGFALHURGKKHVFETONIKGLLLEADYTTKATANLYGLNLNRYF 179
DB 359 TYTLTPQLTMDAGPALVQSRKGSFTEKXNQNLQFSE----AVAYISALQFNRYF 411

RESULT 6
AH0334

probable long-chain fatty acid transport protein fadL [imported] - Yersinia pestis (s
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0334
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92983.1; PID:g15980722; GSPDB:GN00175
C:Genetics:
A:Gene: fadL
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 38.9%; Score 372.5; DB 2; Length 423;
Best Local Similarity 44.6%; Pred. No. 9.4e-27;
Matches 79; Conservative 23; Mismatches 70; Indels 5; Gaps 3;

QY 4 GVLGPIYKSLTLKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGYQDGSYGVE 63
DB 251 GGLGKTVPGSLTLNLPAYWESGVNKPQWAIHYSMAVTTWSSFKELKATASNGD--V 308
QY 64 APTKKEEYKDNRFALGTTYSLNDALTLRAGLAYDKAAST-HLSASIPDTRMYSLCA 122
DB 309 LFDKIEGFADVAIRALGTTYYDDNNWTRTGTAFDPSIPAGNRSISIPDQRFWLSAGT 368
QY 123 TYKFTPNLSVDVGFHLRGKKHFEVETONIKGLLLVEADYTTKATANLYCLNLYRF 179
DB 369 TYAFNKNASVDVGIAVMKCONVITEKTPAPSNVTVE--FNSKGSAMLYGVNFNYTF 423

RESULT 7
AB0805
long-chain fatty acid transport protein precursor STY2623 [imported] - *Salmonella enteri*
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0805
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0805
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07623.1; PID:g16503614; GSPDB:GN00176
C:Genetics:
A:Gene: STY2623
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 35.3%; Score 338; DB 2; Length 437;
Best Local Similarity 42.9%; Pred. No. 1.5e-23;
Matches 73; Conservative 26; Mismatches 55; Indels 16; Gaps 5;

QY 13 GSLTLKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGYQDGSYGVEAPTKKEEK 72
DB 281 GYLTLNLPWNEVSGVNRVAPQWAIHYSLAYTSWSQFQELKAK--STAGDTLPEKHEGPK 338
QY 73 DNSRFAIGTTYSLNDALTLRAGLAYDKA-ASKTHLSASIPDTRMYSLGATYKFTPNLS 131
DB 339 DAYRIALGTTYYDDNNWTRTGTAFDPSVPAQNRSISIPDQRFWLSAGTTVAFNKNKAS 398
QY 132 VDVGFAHLRGKKHFEVETONIKGLLLVEADY--TTKATANLYCLNLYRF 179
DB 399 VDVGVSVMHG-----QSVK---INEGPYQFSECKAWLFGTGFNYAF 437

RESULT 8
B82250
long-chain fatty acid transport protein VC1042 [imported] - *Vibrio cholerae* (strain N169

C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82250
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;
Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
I. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: B82250
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <HEI>
A:Cross-references: GB:AE004185; GB:AE003852; NID:g655497; PIDN:AAF94201.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1042
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 34.9%; Score 334; DB 2; Length 432;
Best Local Similarity 41.3%; Pred. No. 3.6e-23;
Matches 71; Conservative 30; Mismatches 63; Indels 8; Gaps 4;

QY 13 GSUTLKPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGYQDGSYGVEAPTKKEEK 72
DB 264 GMSVTLPAELASFLQNDQWAIHASINWTDSSFKELTAVFEKSD--LIKSENWE 320
QY 73 DNSRFAIGTTYSLNDALTLRAGLAYD-KAASKTHLSASIPDTRMYSLGATYKFTPNLS 131
DB 321 DNYRFALGTTQYDADKALRAGVAYDTSVDDKNRTATIPETDRTWVSGSVVATPQLT 380
QY 132 VDVGFAHLRGKKHFEVETONI--KGLLLVEADYTTKATAN--LYGLNLYRF 179
DB 381 LDAGFTYIFAKDATTINEPRDASDQTAAGGFTGNVGNWLLIGVQANTRF 432

RESULT 9
F65007
long-chain fatty acid transport protein fadL precursor - *Escherichia coli* (strain K-1
N:Alternate names: FLP protein
C:Species: *Escherichia coli*
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: F65007; A39126; S02829
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.,
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F65007
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-448 <BLAT>
A:Cross-references: GB:AE000323; GB:U00096; NID:g1788684; PIDN:AAC75404.1; PID:g17886
A:Experimental source: strain K-12, substrain MG1655
R:Black, P.N.
J. Bacteriol. 173, 435-442, 1991
A:Title: Primary sequence of the *Escherichia coli* fadL gene encoding an outer membran
A:Reference number: A39126; MUID:91100327; PMID:1987139
A:Accession: A39126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177, 'V', 179-287, 'R', 289-448 <BLA>
A:Cross-references: GB:M60607; GB:M37714; NID:g145909; PIDN:AAA64433.1; PID:g145910
R:Said, B.; Ghosh, C.R.; Vu, L.; Nunn, W.D.
Mol. Microbiol. 2, 363-370, 1988
A:Title: Nucleotide sequencing and expression of the fadL gene involved in long-chain
A:Reference number: S02829; MUID:88288050; PMID:2840553
A:Accession: S02829
A:Molecule type: DNA
A:Residues: 66-177, 'R', 179, 'P', 181-287, 'R', 289-448 <SAI>
A:Cross-references: EMBL:X00552; NID:g41371; PIDN:CAA68630.1; PID:g41372
A:Note: the authors translated the codon CTG for residue 162 as Glu

C;Genetics:

A;Gene: fadL

C;Superfamily: long-chain fatty acid transport protein fadL

Query Match 34.7%; Score 332; DB 2; Length 448;
Best Local Similarity 42.9%; Pred. No. 5.7e-23;
Matches 73; Conservative 25; Mismatches 56; Indels 16; Gaps 5;

Qy 13 GSLTLKLPAYWELSGFHQLTDOMAIHYSKYKTEWSRKFELRGKYODGSGYEAFTKKEEK 72

Db 292 GYLTLNLPENMEVSGYRNVDPMQAIHSLAYTSWSQFQL--KATSTSGDTLFOKHEGFK 349

Qy 73 DNSRFAIGTCTTYSNLDAITLRAGLAYDKA-ASKTHLSASIPDTRMYSIGATYKFTPNLS 131

Db 350 DAYRIALGTTTYYDDNMTFTGIAFDSPVPAQNRKSISIPDQRFWLSAGTYYAFNKDAS 409

Qy 132 VDVGFAHLRGKKHVFETONIKGLLLVEADY--TTKATANLYGLNLNRYF 179

Db 410 VDVGVSVMHG-----QSVK---INEGPYQFSEGGKAWLFGTNFYAF 448

RESULT 10

C91032

Long-chain fatty acid transport protein FadL Ecs3227 [imported] - Escherichia coli (stra

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C;Accession: C91032

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C91032

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-448 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA036650.1; PID:gl3362697; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RMD 0509952

C;Genetics:

A;Gene: Ecs3227

C;Superfamily: long-chain fatty acid transport protein fadL

Query Match 34.7%; Score 332; DB 2; Length 448;

Best Local Similarity 42.9%; Pred. No. 5.7e-23;

Matches 73; Conservative 25; Mismatches 56; Indels 16; Gaps 5;

Qy 13 GSLTLKLPAYWELSGFHQLTDOMAIHYSKYKTEWSRKFELRGKYODGSGYEAFTKKEEK 72

Db 292 GYLTLNLPENMEVSGYRNVDPMQAIHSLAYTSWSQFQL--KATSTSGDTLFOKHEGFK 349

Qy 73 DNSRFAIGTCTTYSNLDAITLRAGLAYDKA-ASKTHLSASIPDTRMYSIGATYKFTPNLS 131

Db 350 DAYRIALGTTTYYDDNMTFTGIAFDSPVPAQNRKSISIPDQRFWLSAGTYYAFNKDAS 409

Qy 132 VDVGFAHLRGKKHVFETONIKGLLLVEADY--TTKATANLYGLNLNRYF 179

Db 410 VDVGVSVMHG-----QSVK---INEGPYQFSEGGKAWLFGTNFYAF 448

RESULT 11

D85876

Long-chain fatty acid transport protein FadL - Escherichia coli (strain O157:H7, substra

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C;Accession: D85876

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: D85876

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-448 <STO>

A;Cross-references: GB:AE005174; NID:gl2516711; PIDN:AAG57472.1; GSPDB:GN00145; UWGP:

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: fadL

C;Superfamily: long-chain fatty acid transport protein fadL

Query Match 34.7%; Score 332; DB 2; Length 448;

Best Local Similarity 42.9%; Pred. No. 5.7e-23;

Matches 73; Conservative 25; Mismatches 56; Indels 16; Gaps 5;

Qy 13 GSLTLKLPAYWELSGFHQLTDOMAIHYSKYKTEWSRKFELRGKYODGSGYEAFTKKEEK 72

Db 292 GYLTLNLPENMEVSGYRNVDPMQAIHSLAYTSWSQFQL--KATSTSGDTLFOKHEGFK 349

Qy 73 DNSRFAIGTCTTYSNLDAITLRAGLAYDKA-ASKTHLSASIPDTRMYSIGATYKFTPNLS 131

Db 350 DAYRIALGTTTYYDDNMTFTGIAFDSPVPAQNRKSISIPDQRFWLSAGTYYAFNKDAS 409

Qy 132 VDVGFAHLRGKKHVFETONIKGLLLVEADY--TTKATANLYGLNLNRYF 179

Db 410 VDVGVSVMHG-----QSVK---INEGPYQFSEGGKAWLFGTNFYAF 448

RESULT 12

B82248

Long-chain fatty acid transport protein VC1043 [imported] - Vibrio cholerae (strain N

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: B82248

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

Cardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, L.; Sellers

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82248

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-428 <HEI>

A;Cross-references: GB:AE004186; GB:AE003852; NID:g9655507; PIDN:AAF94202.1; GSPDB:GN

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1043

A;Map position: 1

C;Superfamily: long-chain fatty acid transport protein fadL

Query Match 28.0%; Score 267.5; DB 2; Length 428;

Best Local Similarity 39.3%; Pred. No. 5e-17;

Matches 68; Conservative 21; Mismatches 71; Indels 13; Gaps 5;

Qy 15 LTLKLPAYWELSGFHQLTDOMAIHYSKYKTEWSRKF--ELRGKYODGSGYEAFTKKE-- 69

Db 261 ITLPLPDAIESGFGHKIENTOFVHYSVOWICGSAFDKIEFRLNLTSSSSVCAITKGSYD 320

Qy 70 ---EYKDNRSFALCTTYSNLDAITLRAGLAYDKAASKTHLSASIPDTRMYSIGATYK 126

Db 321 KMYEQDQWHYSIGCTTYVNSDWTLRAGYMYDTSIQDSRTSISVPDSDROWLSAGTYNI 380

Qy 127 TPNLSVDVGFAHLRGKKHVFETONIKGLLLVEADYTTKATANLYGLNLNRYF 179

Db 381 DDKSNVDGFGTYLMG--KDAVSENDAGSKLTA---TTHADAILLGLQYSRTF 428

RESULT 13

E82729

Outer membrane protein XF1053 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C;Accession: E82729

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E87229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>

A:Cross-references: GB:AE003942; GB:AE003849; NID:99105990; PIDN:AAF83863.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Garret, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V
A:Reference number: A59328
A:Contents: annotation
A:Genetics:
A:Gene: XF1053

Query Match 21.6%; Score 207; DB 2; Length 447;
Best Local Similarity 31.8%; Pred. No. 2,1e-11;
Matches 56; Conservative 31; Mismatches 79; Indels 10; Gaps 7;
QY 7 GPYIG-KGSLTKLPAYWELSGFHQLTQDAWHYKYKTEWSRFRKELRGKYQDGSYGAEF 65
DB 279 GTFTNTIKSTIRLPASATTSFTHAIDRWSIMAEVTRTAMSKPKDKITIHFASSPPNNAL 338
QY 66 TKKEEYKDNRRFAIGTYSNLDAITLRAGLAYDKA-ASKTHLSASIPDTRMWSYIGATY 124
DB 339 --NFSYRDTTFVAGTYEFSMDTLRGGLGYDQTPPTAEYRSVRVPDNNRTWSLGMTW 396
QY 125 KFTPNLSVDVGFALHRGKKHFEVETQ-NIKGLLVEADYTTKATANLYGLNLNRYF 179
DB 397 KPSQOTDNYGAYHL--FIGHTTQINATASTLV-GDY--NINTDLLAISNIHF 447

RESULT 14
E87508
outer membrane protein, probable CC2094 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87508
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87508
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <STO>
A:Cross-references: GB:AE005673; NID:913423577; PIDN:AAK24065.1; GSPDB:GN00148
A:Genetics:
A:Gene: CC2094

Query Match 20.3%; Score 194.5; DB 2; Length 432;
Best Local Similarity 29.2%; Pred. No. 2.9e-10;
Matches 52; Conservative 26; Mismatches 77; Indels 23; Gaps 5;
QY 3 NGVLGP-----YTKGSLTKLPAYWELSGFHQLTQDAWHYKYKTEWSRFRKELRGKY 56
DB 253 SGLLAPAAASNETVDGQAKITLPWIANLGARWAVNDQWTLNGSVSRVGVWSEFDAIRVSF 312
QY 57 QDGSYGAEFYKKEYKNSRFAIGTYSNLDAITLRAGLAYDKAAS-KTHLSASIPDTR 115
DB 313 AGGGS----TSPQDYKDVTYAGVDYQASPRLTLRAGVQVDPTPTPEIGRTARVPDGR 368

QY 116 MWYSIGATYKFTPNLSVDVGFALHRGKKHFEVETQNIKGLLVEADYT-TRATANLYG 172
DB 369 MMYATGATWAATETLQDDAALSYAFDKSQINRT-----DVTATSTSVRLRG 415

RESULT 15
A83073
probable outer membrane protein PA4589 [imported] - Pseudomonas aeruginosa (strain P
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83073
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of pseudomonas aeruginosa PA01, an opportunistic p
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <STO>
A:Cross-references: GB:AE004872; GB:AE004091; NID:9950829; PIDN:AAG07977.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4589

Query Match 19.9%; Score 190.5; DB 2; Length 463;
Best Local Similarity 30.6%; Pred. No. 7.5e-10;
Matches 55; Conservative 30; Mismatches 66; Indels 29; Gaps 7;
QY 14 SLTKLPAYWELSGFHQLTQDAWHYKYKTEWSRFRKELRGKYQDGSYGAEFTKKEEYKD 73
DB 299 SARLDIPAYASLDVHWFNDRLSLGASATWTWSSFQDLTLK---SHGNTIVSIPYTYKN 355
QY 74 NSRFAIGTYSNLDAITLRAGLAYDKAASKTHLSA---SIPDTRMWSYIGATYKF--TP 128
DB 356 TWTLAGVGDKYKVTQDWTMRAGVAYDQ--TPTHNATRDPRIPDGRYFASLGAGYRFQSM 413
QY 129 NLSVDVGFALHRGKKHFEVETQNIK-----CLLVEADYTTKATANLYGLNLNRYF 179
DB 414 ELSIDAAYS-----ROFVKVEVPLKTVNQDRLGGRL----DGRATSKGVFSLSATYDF 463

RESULT 16
H83484
probable outer membrane protein PA1288 [imported] - Pseudomonas aeruginosa (strain P
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83484
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83484
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <STO>
A:Cross-references: GB:AE004558; GB:AE004091; NID:99947217; PIDN:AAG04677.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1288

Query Match 17.5%; Score 167; DB 2; Length 424;
Best Local Similarity 27.7%; Pred. No. 1e-07;
Matches 52; Conservative 34; Mismatches 80; Indels 22; Gaps 7;
QY 4 GVLGPYTKGSLTKLPAYWELSGFHQLTQDAWHYKYKTEWSRFRKELRGKYQ----- 57
DB 247 GLNGKYDAK--LDITLPESVDTSITHKFDKDKWTGVLGAVWTRWSRLEKIEVRNSGVPAUG 304

QY 58 DGSYEAFPTKKEEYKDNSRFAIGTTYSUNDALTLRAGLAYDKA-ASKTHLSASIPDTRM 116
Db 305 QALGFNTIGEDLNLWDTWSFVSGYQATPEWLRTGFAVEPSPTSNEEDNRVIRPVGDRK 364
QY 117 WYSICATYKFTPNLSVDVGFAGHLRKKKHFEVETO---NIKGLLLVEADYTTKATNLYGL 173
Db 365 VFTVGAGSPNQDLTVDAAYL-----WETTAGVNOEGSALOPA-YSAKYDNSAHGL 416
QY 174 --NLNYRF 179
Db 417 TAQTYRF 424
RESULT 17
A81239
outer membrane protein p1, probable NMB0088 [imported] - Neisseria meningitidis (strain
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: A81239
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: A81239
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <TET>
A:Cross-references: GB:AE002368; GB:AE002098; NID:97225303; PIDN:AAF40551.1; PID:g722530
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0088
Query Match 15.18; Score 144.5; DB 2; Length 466;
Best Local Similarity 23.48; Pred. No. 1.4e-05;
Matches 45; Conservative 38; Mismatches 70; Indels 39; Gaps 7;
QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSKYKTEWSRF-----KE---LRGKYODGS 60
Db 289 KARVKIVTPESLSVHGMYKVSADKADLFGDVTWTRHSRFDKAEVFEKTVVKCK----- 343
QY 61 GYEAFPTKKEEYKDNSRFAIGTTYSUNDALTLRAGLAYDKAASKT--HLSASIPDTRMWY 118
Db 344 -SDRTTTPNNRNTYKVGFGSGYQISEPLQLRAGIAFDKSPVRNADYRMNSLPGDNRIWF 402
QY 119 SIGATYKFTPNLSVDVGFAGHLRKKKHFEVETO-----NIKGLLLVEADYTTKAT 167
Db 403 SAGMKYHIGKNHVVDAAAYTHI-----HINDTSYRTAKASGNDVDSKG-----ASSARFKNH 453
QY 168 ANLYGLNLNYRF 179
Db 454 ADIIGLQYTYKF 465
RESULT 18
G83424
hypothetical protein PA1764 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83424
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83424
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-532 <STO>
A:Cross-references: GB:AE004602; GB:AE004091; NID:g9947739; PIDN:AGC05153.1; GSPDB:GN001

A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA1764
Query Match 15.08; Score 143.5; DB 2; Length 532;
Best Local Similarity 30.08; Pred. No. 2e-05;
Matches 48; Conservative 21; Mismatches 60; Indels 31; Gaps 9;
QY 34 QWAIHYSVK---YTEMSREF-----ELR-----GKYODG--SGYEAFPTKKEEYKDNSRFAI 79
Db 363 KWQFNFDLKNWGYSDWNEFKIEFDRELDLLRIAYLDSKDATSHITLDRGYRDTWSWAM 422
QY 80 GTTYSUNDALTLRAGLAY-DKAASKTHLSASIPDTRMWYSIGATYKFTPNLSVDVGFAGH 138
Db 423 GVQYDVNDRLQLRAGYEVRPSAIPKGQADILVPIGDANLYGLGLGYQWDKTDVIDVGF-- 480
QY 139 LRGKKKHFEVETOIKGLLLVEADYTTKATNLYGL-NLNY 177
Db 481 ----NYFVTKQSVK-----AD--NSCNLCTGLDNLVY 507
RESULT 19
H82011
probable outer membrane protein NMA0178 [imported] - Neisseria meningitidis (strain 2
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: H82011
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: H82011
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-464 <PAR>
A:Cross-references: GB:AL162752; GB:AL157959; NID:97378778; PIDN:CAB83493.1; PID:g737
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMA0178
Query Match 14.98; Score 142.5; DB 2; Length 464;
Best Local Similarity 23.48; Pred. No. 2.1e-05;
Matches 45; Conservative 37; Mismatches 71; Indels 39; Gaps 7;
QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSKYKTEWSRF-----KE---LRGKYODGS 60
Db 287 KARVKIVTPESLSVHGMYKVSADKADLFGDVTWTRHSRFDKAEVFEKTVVKCK----- 341
QY 61 GYEAFPTKKEEYKDNSRFAIGTTYSUNDALTLRAGLAYDKAASKT--HLSASIPDTRMWY 118
Db 342 -SDRTTTPNNRNTYKVGFGSGYQISEPLQLRAGIAFDKSPVRNADYRMNSLPGDNRIWF 400
QY 119 SIGATYKFTPNLSVDVGFAGHLRKKKHFEVETO-----NIKGLLLVEADYTTKAT 167
Db 401 SAGMKYHIGKNHVVDAAAYTHI-----HINDTSYRTAKASGNDVDSKG-----ASSARFKNH 451
QY 168 ANLYGLNLNYRF 179
Db 452 ADIIGLQYTYKF 463
RESULT 20
C70385
hypothetical protein aq_985 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70385
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320

Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: C71889
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <ARN>
A:Cross-references: GB:AE001508; GB:AE001439; NID:g4155338; PIDN:AAD06357.1; PID:g415534
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0777

Query Match 10.7%; Score 102; DB 2; Length 587;
Best Local Similarity 25.5%; Pred. No. 0.16;
Matches 42; Conservative 23; Mismatches 62; Indels 38; Gaps 8;

Qy 6 LGPYIG----KGSUTL--KLPAWELSGFHQ-LTDQWAIHYSKYTWSRFKEL----- 52
Db 396 LGPSLGSVLTKSUNINVSUPTLSLAYAHQFKDRLURGVGFERTFWSQGNKFLVTPDF 455
Qy 53 -RGYQDGSVEAFTKKEE-----YKDNSRFAIGTYSLNDALTLR 92
Db 456 ANATYKGLSGTVASLDSLETLKKWGLANFKSVNMWAGWRDTNTRFLGVTY-MCKSLRLM 514
Qy 93 AGLAYDKAASKTHLSASIPDTRMWSIGATYKFTPNLSVDVGF 137
Db 515 GAIDYDQAPSPQD-AIGIPDSNGYTVAFGTQYNFR---GFDLGVA 555

RESULT 25
D83081
probable outer membrane receptor for iron transport PA4514 [Imported] - Pseudomonas aeru
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 14-Sep-2001
C:Accession: D83081
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83081
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-753 <STO>
A:Cross-references: GB:AE004865; GB:AE004091; NID:g9950752; PIDN:AAG07902.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4514
C:Superfamily: ferrichrome-iron receptor 1; tonB-dependent receptor amlno-terminal homol

Query Match 10.4%; Score 100; DB 2; Length 753;
Best Local Similarity 27.5%; Pred. No. 0.33;
Matches 33; Conservative 20; Mismatches 53; Indels 14; Gaps 6;

Qy 27 GFH-QLTDQWAIHYSKYTWSRFKELRGYQDGSVEATKKEEYKDNSRFAIGTYSYL 85
Db 618 GFNGKLEKWKVGGYTYLD-SEAKSTVKSDEGN-----KMPOTAQNRFTLWTTTYDL 669
Qy 86 NDALTLAGLAY-DKAASKTHLSASIPDTRMW-YSIGATYKFTPNLSVDVGFALRGKK 143
Db 670 LQNFITGGTYYVDKQGNANTSTYIPS---YWRDAMASYKSKNVLDLQNLVNTDKR 726

RESULT 26
B81038
TonB-dependent receptor NMB1829 [Imported] - Neisseria meningitidis (strain MC58 serogrp
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81038
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rl, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000.
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis sdrgroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <TET>
A:Cross-references: GB:AE002532; GB:AE002098; NID:g7227078; PIDN:AAF42164.1; PID:g7222
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1829

Query Match 10.1%; Score 96.5; DB 2; Length 708;
Best Local Similarity 24.4%; Pred. No. 0.65;
Matches 44; Conservative 25; Mismatches 70; Indels 41; Gaps 9;

Qy 6 LGPYIGKGSUTLKLPAWELSGFHQ-LTDQWAIHYSKYTWSRFKELRGYQDGSVEAF 65
Db 564 LGRVMEGVET-----EISG--AMTPKQIHAGYSYLH-SQIKTASNSRDEG----IF 609
Qy 66 TKREYKDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSIGATYK 125
Db 610 LLMPKUISAN---LWTTYQVTSLTGIGGV---NAMSGITSSAGIHAGGYATPDMAAYR 662
Qy 126 FTPNLSVDVGFALRGKKHFVETQNIKGLLLVADYTTKATANLYG-----LNLNRYF 179
Db 663 FTPKCLKQINADNI--FNRHY-----ARVSESTFNIPGSRSLTANLRYSF 708

RESULT 27
AD1209
N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28) lmo1076 - Listeria monocytogenes (st
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 31-Dec-2001
C:Accession: AD1209
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99154.1; PID:g16410478; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1076
C:Keywords: hydrolase

Query Match 9.8%; Score 94; DB 2; Length 572;
Best Local Similarity 26.1%; Pred. No. 0.84;
Matches 57; Conservative 30; Mismatches 67; Indels 64; Gaps 12;

Qy 4 GVLGPYIGKGSUTLKLPAWELSGFHQ-LTDQWAIHYSKYTWSRFKELR-GKYQDGSY 62
Db 140 GIKGSYNGK-SVPMKWTWEYSDSKGWQYQINANFAPKYPKSHKESLEDNAKLNKNGPSWDSYY 198;
Qy 63 EAFTKK--EBYKD-----NSRFAIGTYS--LN-----D 87
Db 199 KGAWRENAKTYKDATAWLQGRYATDNTYASKNLTLISSNLTQYDLYDTLKQKNVSED 258
Qy 88 ALTLRA-----GLAYDKAASKTHLSASIPDTR-----MWSIGATYKFTPNLSVD 133
Db 259 AKVVKADGCHGVSGIYNTSAASAKKLSTGAPYNNKDKYLKKEGTTSRGTWQVFSNNKV- 317
Qy 134 VGFAHLRGKKHVF---ETONIKGLLLVADYTTKATA 168
Db 318 IGWV---DKRAFVYKPKATNKTNLN-----TGKITA 346

```

RESULT 28
D69426
surface layer protein B (slqB-2) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: D69426
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
:; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Sprlgs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: D69426
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2425 <KLE>
A:Cross-references: GB:AE001006; GB:AE000782; NID:g2689329; PIDN:AAB89834.1; PID:g264915

Query Match          9.7%; Score 92.5; DB 2; Length 2425;
Best Local Similarity 25.2%; Pred. No. 7.3;
Matches 36; Conservative 17; Mismatches 43; Indels 47; Gaps 8;

QY 3 NGVLGPYIGKSLTLKLPAY-----ELSGFHQLTDQWAIHYSKYKTEWSRFKELRGKY 56
DB 998 NIIGPYLGGN-----YHDIYAGEDVAGDGLGDTLLPYN-----GGMI 1037

QY 57 QGSGYEAFTK-KEEYKNSRFAT-----GTTYSLNDALTFRAGLAYDKAASKTHLSASIP 111
DB 1038 QNGGDYHPLTNPDPFRAPAIYVSPVEGRYSAN-----YVLEVYSPD-----P 1083

QY 112 DTRDWYSI--CATYKFTPNLSV 132
DB 1084 DVDRWYLSLNSGANVTETPNNTTI I106

RESULT 29
B83592
hypothetical protein PA0434 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83592
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
:; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83592
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-730 <STO>
A:Cross-references: GB:AE004480; GB:AE004091; NID:g9946284; PIDN:AA03823.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0434

Query Match          9.6%; Score 91.5; DB 2; Length 730;
Best Local Similarity 25.4%; Pred. No. 2;
Matches 34; Conservative 16; Mismatches 45; Indels 39; Gaps 7;

QY 31 LTDQWAIHYSKYKTEWSRFKELR-----GKYQDSGYEAFTK-----EYKD 73
DB 377 LASDWALGFDYSINKQTRFPPTASGPGFTVDPASPEPGHFYDLPGRPCHRKDRSNEVRT 436

QY 74 NSRFAIGTYSLNDALTFRAGLAYDKAASKTHLSAS-----IPDPTD-----RWYSI-- 120
DB 437 SALFA-ENRLGTDLSLVTGLRYD-----HLDLDVRNHRVPDKDNFAHFERRDVTG 489

QY 121 --CATYKFTPNLSV 132

```

```

DB 490 RAGLVYQFTPHANV 503
| :|:|:|:|

RESULT 30
D90755
outer membrane protein la ECS1012 [imported] - Escherichia coli (strain O157:H7, subs
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: D90755
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D90755
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA034435.1; PID:g13360471; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS1012
C:Superfamily: outer membrane protein phoE

Query Match          9.5%; Score 90.5; DB 2; Length 362;
Best Local Similarity 25.9%; Pred. No. 1;
Matches 37; Conservative 22; Mismatches 59; Indels 25; Gaps 6;

QY 4 GVLGPYIGKSLTLKLPAYWELSGFHQLTDQW-----IHYSKYKTEWSRFKELRG 54
DB 208 GIVGAY-GAADRNLQEA--QLLNGKKAQWATGLKYDANNIYLAANGETRNATPTN 264

QY 55 KYQDSGYEAFTKKEEYKNSRFATGTTYSLNDALTFRAGLAYDKAASKTHLSASIPDPTD 114
DB 265 KFTNISGSEFANKQDVLVLAQYQDFG-----LRPSIATYKSKAKD--VEGIGDVD 312

QY 115 RM-WYSIGATYKFTPNLSVDVGF 136
DB 313 LVNYFEVGATYFENKMWSTYVDY 335

RESULT 31
B85619
outer membrane protein la (Ia,b,F) [imported] - Escherichia coli (strain O157:H7, sub
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85619
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <STO>
A:Cross-references: GB:AE005174; NID:g12514104; PIDN:AA055414.1; GSPDB:GN00145; UMG
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ompF
C:Superfamily: outer membrane protein phoE

Query Match          9.5%; Score 90.5; DB 2; Length 362;
Best Local Similarity 25.9%; Pred. No. 1;
Matches 37; Conservative 22; Mismatches 59; Indels 25; Gaps 6;

QY 4 GVLGPYIGKSLTLKLPAYWELSGFHQLTDQW-----IHYSKYKTEWSRFKELRG 54
DB 208 GIVGAY-GAADRNLQEA--QLLNGKKAQWATGLKYDANNIYLAANGETRNATPTN 264

QY 55 KYQDSGYEAFTKKEEYKNSRFATGTTYSLNDALTFRAGLAYDKAASKTHLSASIPDPTD 114
DB 265 KFTNISGSEFANKQDVLVLAQYQDFG-----LRPSIATYKSKAKD--VEGIGDVD 312

```



```
RESULT 38
A82358
vitamin B12 receptor VC0156 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82358
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Karchison, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: A82358
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-611 <HEI>
A:Cross-references: GB:AE004105; GB:AE003852; NID:99654551; PIDN:AAF93332.1.; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0156
A:Map position: 1
C:Superfamily: vitamin B12 receptor; tonB-dependent receptor amino-terminal homology; to
Query Match 8.9%; Score 85.5; DB 2; Length 611;
Best Local Similarity 19.9%; Pred. No. 5.6;
Matches 51; Conservative 23; Mismatches 97; Indels 85; Gaps 8;
QY 1 QHNGVLGPYIGKSLTLKLPAY--WELSGFHQDQWAIHYSKYTE-WSRFKELRGK-- 55
DB 189 QMNVALGTESDKGNVRPVGNDGRHGR--SDNALLGVVHGFDESWSLFANARAYEN 246
QY 56 -YQDGSYEAFTKKEEYKDNSRFAIGTYS----- 84
DB 247 IQYDINSYGRDYKEAEKDDLSFTIGTOYQSERWVSELQLTQKQSWDYTSKGYSDT 306
QY 85 -----LNDALTLRAGLAY-----DKAASKTH----- 105
DB 307 SDNLEQRNIOWINRYLVNDVMTFAGGVWDRDESYIDKADFEORSNTAFAVVAEWMQ 366
QY 106 --LSASIPDTRMYY-----SIGATYKFTPNLSVDVGVGFALHKGKKHFEVTONIKGLLL 157
DB 367 WLLASLRFDNDQYGSQTHNIALGQYFPEFCVKASYGSFAKAPNLYQQYDPSYNNV 426
QY 158 VEADYTKATANLYGL 173
DB 427 LQPEDADSALSFYGL 442
RESULT 39
G97053
penicillin-binding protein 2 [Imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: G97053
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-916 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79218.1; PID:g15024172; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1246
Query Match 8.9%; Score 85.5; DB 2; Length 916;
Best Local Similarity 25.6%; Pred. No. 9.4;
Matches 31; Conservative 18; Mismatches 33; Indels 39; Gaps 5;
QY 28 FHQLTDQWAIHYSY-KYTEWSRFKELR-GKYQDGSYEAFTK-----KEEYKDNS 75
```

```
DB 606 FGQVNTYTIKNSFADEGQWIAEKLOEGKYLKGSVNLYTSDSKSLTLKSIKENI 665
QY 76 RFAL-----GTTYSLNDALTLRAGLAYDKAASKTHLSASI 110
DB 666 RTAISTGKTDRNKTEELVTLIDTLDPLKYGKTSKSDVSALASGIVYDEADEYT--QATV 723
QY 111 P 111
DB 724 P 724
RESULT 40
F84148
hypothetical protein BH3990 [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F84148
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F84148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07709.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3990
Query Match 8.9%; Score 85; DB 2; Length 221;
Best Local Similarity 24.1%; Pred. No. 1.7;
Matches 34; Conservative 20; Mismatches 51; Indels 36; Gaps 6;
QY 26 SGFHQLTDQWAIHYSKYTEWSRFKELRGKYQ-----DSGGYEAF 65
DB 67 SGFSTMSTSSSSSKAQHKYNSKS-FSHAKGGYSVIDVKRSKPGFWRSLYQHRYDGYGTAY 125
QY 66 TKKEEYKDNSRFA-----IGTYSLN-----DALTLRAGLAYDKA--ASKTHLSASIP 111
DB 126 TAKKKYKPSNSAKVATKICIPNVNLSRNSDWSLLTSSRTVAIDSAKYSRKYAQANPD 185
QY 112 DTRMWSY-IGATYKFTPNLS 131
DB 186 DLRIQWLSNLAVSPKVESKLS 206
Search completed: May 12, 2003, 09:41:39
Job time.: 20 secs
```


GenCore version 5.1.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:38:34 ; Search time 13 Seconds
(without alignments)
571.097 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 QHNGVLGPYICKGSLTKLP.....ADYTTKATANLYGLNLNRYR 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	549	57.4	459	OM11_HAEIN	P43838 haemophilus
2	549	57.4	459	OM12_HAEIN	P10641 haemophilus
3	539.5	56.4	428	OM47_PASMU	P80603 pasteurella
4	332	34.7	448	FADL_ECOLI	P10384 escherichia
5	90.5	9.5	362	OMPD_SALTY	P37592 salmonella
6	90.5	9.5	875	FODD_ECOLI	P46009 escherichia
7	88.5	9.2	522	MSAB_NEIMB	Q9K1N8 n peptide m
8	88	9.2	479	Y865_TREPA	O83837 treponema p
9	87.5	9.1	810	HPUB_NEIMC	P96949 neisseria m
10	86.5	9.0	522	MSAB_NEIGO	P14930 n peptide m
11	85.5	8.9	362	OMPF_ECOLI	P02931 escherichia
12	84	8.8	810	HPUB_NEIMA	Q91W42 neisseria m
13	83.5	8.7	362	ARYH_RAT	Q02589 rattus norv
14	83	8.7	937	CS32_ECOLI	P15484 escherichia
15	82.5	8.6	522	MSAB_NEIMA	Q9JW88 n peptide m
16	81.5	8.5	362	ARYH_MOUSE	P54923 mus musculus
17	80.5	8.4	386	ALR_CLOPE	Q8XM22 clostridium
18	80	8.4	265	ARC2_PHAVU	P19330 phaseolus v
19	80	8.4	365	NMPC_ECOLI	P21420 escherichia
20	80	8.4	698	TREH_HUMAN	P02787 homo sapien
21	80	8.4	968	CTDI_HUMAN	O60716 homo sapien
22	79.5	8.3	265	ARCI_PHAVU	P19329 phaseolus v
23	79	8.3	911	CTDI_MOUSE	P30999 mus musculus
24	78.5	8.2	185	PAGC_SALTY	P23988 salmonella
25	78.5	8.2	345	YE6A_HAEIN	O86241 haemophilus
26	78.5	8.2	720	FPTA_PSEAE	P42312 pseudomonas
27	78.5	8.2	1715	NX2A_RAT	O63374 rattus norv
28	78	8.2	178	AIL_YEREN	P16454 versinia en
29	77	8.0	252	FENR_BUCAI	P57641 buchiera ap
30	77	8.0	365	PORI_BPPA2	P07238 bacterioph
31	76.5	8.0	774	FECA_ECOLI	P13036 escherichia
32	76	7.9	289	PORI_RHOBL	P39767 rhodopseudo
33	75.5	7.9	357	ARYH_HUMAN	P54922 homo sapien

ALIGNMENTS

RESULT 1

ID	OM11_HAEIN	STANDARD;	PRT;	459 AA.
AC	P43838;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Outer membrane protein P1 precursor (OMP P1).			
GN	OMP1 OR H10401.			
OS	Haemophilus influenzae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
OC	Haemophilus.			
OX	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Rd / KW20 / ATCC 51907;			
RX	MEDLINE=95350630; PubMed=7542800;			
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,			
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,			
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,			
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,			
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,			
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,			
RA	Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,			
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,			
RT	Venter J.C.;			
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae			
RL	Rd.;"			
RN	Science 269:496-512(1995).			
RP	IDENTIFICATION BY MASS SPECTROMETRY.			
RX	MEDLINE=20137488; PubMed=10675023;			
RA	Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,			
RA	Gray C., Fountoulakis M.;			
RT	"Two-dimensional map of the proteome of Haemophilus influenzae.;"			
RL	Electrophoresis 21:411-429(2000).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.			
CC	-1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@lsb-sib.ch).			
CC	-----			
DR	EMBL: U32723; AAC22060.1; --			
DR	TLGR; H10401; --			
KW	Outer membrane; Signal; Transmembrane; Complete proteome.			
FT	SIGNAL 1 22 BY SIMILARITY.			
FT	CHAIN 23 459 OUTER MEMBRANE PROTEIN P1.			
FT	SEQUENCE 459 AA; 49477 MW; 43BDC98BE5A39366F CRC64;			
SQ				

Query Match 57.4%; Score 549; DB 1; Length 459;
Best Local Similarity 61.6%; Pred. No. 1.1e-43;

RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-38.
 RX MEDLINE-91100327; PubMed-1987139;
 RA Black P.N.;
 RT "Primary sequence of the Escherichia coli fadL gene encoding an outer
 RL membrane protein required for long-chain fatty acid transport.";
 RN J. Bacteriol. 173:435-442(1991).
 RP SEQUENCE FROM N.A.
 RN [2]
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1234-1238(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RN STRAIN-K12;
 RX MEDLINE-97349980; PubMed-9205837;
 RA Yanamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Mikl T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 RA Oshima T., Oyama S., Salto N., Sempel G., Satoh Y., Sivasubramanian S.,
 RA Tadami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yanagata S., Horikuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
 RL - K12 genome corresponding to 50.0-68.8 mln on the linkage map and
 RT analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 RN [4]
 RP SEQUENCE OF 66-448 FROM N.A.
 RN STRAIN-K12;
 RX MEDLINE-88288050; PubMed-2840553;
 RA Said B., Ghosh C.R., Vu L., Nunn W.;
 RT "Nucleotide sequencing and expression of the fadL gene involved in
 RT long-chain fatty acid transport in Escherichia coli.";
 RL Mol. Microbiol. 2:363-370(1988).
 CC -1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
 CC ACROSS THE OUTER MEMBRANE. IT IS A RECEPTOR FOR THE BACTERIOPHAGE
 CC T2. FADL MAY FORM A SPECIFIC CHANNEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC CONTAINS SEVERAL POTENTIAL TRANSMEMBRANE DOMAINS.
 CC -1- INDUCTION: BY LONG-CHAIN FATTY ACIDS. EXPRESSION OF FADL IS UNDER
 CC THE CONTROL OF THE FADL REPRESSOR.
 CC -1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60607; AAA64433.1; -;
 DR EMBL; AE000323; AAC75404.1; -;
 DR EMBL; D90865; BAA16205.1; -;
 DR EMBL; F00552; CAA68630.1; -;
 DR PIR; A39126; A39126.
 DR PIR; S02829; S02829.
 DR EcoGene; EG10280; fadL.
 KW Outer membrane; Lipid transport; Phage recognition; Signal;
 KW Transmembrane; Complete proteome.
 FT SIGNAL
 FT CHAIN 1 27
 FT 28 448 LONG-CHAIN FATTY ACID TRANSPORT PROTEIN.
 FT 178 178 A -> V (IN REF. 1).
 FT CONFLICT 178 180 ARA -> RRP (IN REF. 4).
 FT CONFLICT 288 288 R -> A (IN REF. 2 AND 3).
 SQ SEQUENCE 448 AA; 48857 MW; BJA48EA07210FA88 CRC64;
 Query Match 34.7%; Score 332; DB 1; Length 448;
 Best Local Similarity 42.9%; Pred. No. 1.7e-23;

Matches 73; Conservattive 25; Mismatches 56; Indels 16; Gaps 5;
 QY 13 GSLTLKLPAWELSGFHLTDQWAIHYSKYKTEWSEKELRCKYODGSCYEAFKKKEVK 72
 DB 292 GYLTLNLPMEVSGYNRVDQWAIHYSKYKTEWSEKELRCKYODGSCYEAFKKKEVK 349
 QY 73 DNSRFAIGTYSNDALTLRAGLAYDKA-ASKTHLSASIPDPTDRMYSIGATYKFTPNLS 131
 DB 350 DAYRIALGCTYYDDNWTFTGIAFDDSPVPAQNSISIPDQDRFLWSAGTYYAFNKDAS 409
 QY 132 VDVGFARHGRKKKHVEVQNTKGLLLVREAY--TTKATANLYGLNLNRYF 179
 DB 410 VDVGVSYMHG-----QSVK---INEGYQFSEGRKAWLFGTNPYAF 448
 RESULT 5
 OMPD_SALTY STANDARD; PRT; 362 AA.
 AC P37592;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Outer membrane porin protein ompD precursor.
 GN OMPD OR NMPC OR STM1572.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spleth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RL LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE OF 80-362 FROM N.A.
 RC STRAIN-SL1303;
 RX MEDLINE-95011654; PubMed-7926834;
 RA Hongo E., Morimyo M., Mita K., Machida I., Hama-Inaba H., Tsuji H.,
 RA Ichimura S., Noda Y.;
 RT "The methyl viologen-resistance-encoding gene smva of Salmonella
 RL typhimurium.";
 RL Gene 148:173-174(1994).
 RN [3]
 RP IDENTIFICATION AS OMPD.
 RA Singh S.P., Miller S., Williams Y.U., Rudd K.E., Nikaido H.;
 RL Unpublished observations (FEB-1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- SIMILARITY: BELONGS TO THE OMP/PHOE FAMILY OF PORINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE008768; AAL20490.1; -;
 DR EMBL; D26057; BAA05056.1; -;
 DR HSSP; P02931; 1GPN.
 DR StyGene; SGI0249; ompD.
 DR InterPro; IPR003229; OMP_2.
 DR InterPro; IPR001702; Porin-gram-ve.
 DR Pfam; PF00267; Gram-ve_porins; 2.
 DR PROSITE; PS00576; GRAM_NEG_PORIN; 1.
 KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.

```

Best Local Similarity 22.0%; Pred. No. 1.1;      8;
Matches 44; Conservative 29; Mismatches 64; Indels 63; Gaps

QY 9 YICKGSLTKLPAYWELSGFHQLTDQWALHYSYKYTWSRF-----KELCKYQDGS--- 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 YSTRGYEFTADTAKWKGSGSVLTQGVIOIQPKYTDYNYLAYNKGECAGEYQPADGESS 539
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 -----YEAF---TKKEEYKDNRSFAIGT---TYSLNDALTLFRAG-----LAYDKAAS 102
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 540 TLYLGSCHQSYMGTDTRDRLQNAGENSSVNDISWLSNYSLSRNAQWETORILSFDVSLP 599
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 103 KTHLSASIPOTDRMWYSIGATYFTPNLSVDVGFALHRGKKHIEVETQNTKGLLLVADY 162
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 600 FSHWMRS--DSTSAWRNASARYSQT-----LEAHG 627
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 163 TTKATANLYGL-----NLNY 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 628 QAASTAGLYGTTLLGDNNLGY 647
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

DE Peptide methionine sulfoxide reductase msrA/msrB (EC 1.8.4.6)
 DE [includes: Thioredoxin; Peptide methionine sulfoxide reductase marA
 DE (protein-methionine-S-oxide reductase) (Peptide Met(O) reductase);
 DE Peptide methionine sulfoxide reductase msrB].
 DE MSRB OR PILB OR NMB0044.
 GN
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=MC58 / Serogroup B;
 RC MEDLINE=20175755; PubMed=10710307;
 RX Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.D., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Cittono H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 RT MC58";
 RL Science 287:1809-1815(2000).
 CC -!- FUNCTION: Has an important function as a repair enzyme for
 CC proteins that have been inactivated by oxidation. Catalyzes the
 CC reversible oxidation-reduction of methionine sulfoxide in
 CC proteins to methionine (by similarity).
 CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
 CC protein L-methionine S-oxide + reduced thioredoxin.
 CC -!- DOMAIN: Possesses 2 methionine sulfoxide reductase domains (A/Msra
 CC and B/Msrb) and 1 N-terminal thioredoxin domain. The domain B
 CC exhibits a thioredoxin dependent methionine sulfoxide reductase
 CC activity; The Cys-495 is probably involved in the reduction of
 CC Metso and in formation of the sulfenic acid derivative. The
 CC regeneration of Cys-495 is probably done via formation of a
 CC disulfide bond with Cys-440 followed by its reduction by
 CC thioredoxin.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOREDOXIN
 CC FAMILY.
 CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE MSRA MET
 CC SULFOXIDE REDUCTASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE MSRB MET
 CC SULFOXIDE REDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation

This Swiss entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpost -

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=83090452; PubMed=6294623;
RA Inokuchi K., Mutoh N., Matsuyama S.-I., Mizushima S.;
RT "Primary structure of the ompF gene that codes for a major outer
RT membrane protein of Escherichia coli K-12.";
RL Nucleic Acids Res. 10:6957-6968(1982).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampaio G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
[4]
RN SEQUENCE OF 1-37 FROM N.A.
RP MEDLINE=82139379; PubMed=7037455;
RX Mutoh N., Inokuchi K., Mizushima S.-I.;
RA "Amino acid sequence of the signal peptide of OmpF, a major outer
RT membrane protein of Escherichia coli.";
RL FEBS Lett. 137:171-174(1982).
[5]
RN SEQUENCE OF 23-362.
RX MEDLINE=82256494; PubMed=7049161;
RA Chen R., Kramer C., Schmidmayr W., Chen-Schmeissner U., Henning U.;
RT "Primary structure of major outer-membrane protein I (ompF protein,
RT porin) of Escherichia coli B/r.";
RL Biochem. J. 203:33-43(1982).
[6]
RN SEQUENCE OF 33-63 FROM N.A.
RX MEDLINE=86033642; PubMed=2997131;
RA Nogami T., Mizuno T., Mizushima S.;
RT "Construction of a series of ompF-ompC chimeric genes by in vivo
RT homologous recombination in Escherichia coli and characterization of
RT the translational products.";
RL J. Bacteriol. 164:797-801(1985).
[7]
RN SEQUENCE OF 23-34 AND 39-47.
RP STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robinson K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
[8]
RN SEQUENCE OF 23-27.
RP STRAIN=K12 / W3110;
RX MEDLINE=98291876; PubMed=9629924;
RA Molloy M.P., Herbert B.R., Walsh B.J., Tyler M.I., Traini M.,
RA Sanchez J.-C., Hochstrasser D.F., Williams K.L., Gooley A.A.;
RT "Extraction of membrane proteins by differential solubilization for
RT separation using two-dimensional gel electrophoresis.";

RL Electrophoresis 19:837-844(1998).
[9]
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=92375189; PubMed=1380671;
RA Cowan S.W., Schirmer T., Rummel G., Steiert M., Ghosh R.,
RA Paupit R.A., Jansonius J.N., Rosenbusch J.P.;
RT "Crystal structures explain functional properties of two E. coli
RT porins.";
RL Nature 358:727-733(1992).
[10]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF MUTANT ASP-141.
RX MEDLINE=95024177; PubMed=7524100;
RA Jeanteur D., Schirmer T., Fourel D., Simonet V., Rummel G., Widmer C.,
RA Rosenbusch J.P., Pattus F., Pages J.M.;
RT "Structural and functional alterations of a colicin-resistant mutant
RT of OmpF porin from Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10675-10679(1994).
[11]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=99060042; PubMed=9843370;
RA Phale P.S., Philippsen A., Kiefhaber T., Koebnik R., Phale V.P.,
RA Schirmer T., Rosenbusch J.P.;
RT "Stability of trimeric OmpF porin: the contributions of the latching
RT loop L2.";
RL Biochemistry 37:15663-15670(1998).
[12]
RN SMALL MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE OUTER
CC MEMBRANE. IT IS ALSO RECEPTOR FOR THE BACTERIOPHAGE T2.
CC
CC -1- SUBUNIT: HOMOTRIMER.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC
CC -1- SIMILARITY: BELONGS TO THE OMP/PHOE FAMILY OF PORINS.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: J01655; AAA24244.1; -;
DR EMBL: AE000195; AAC74015.1; -;
DR EMBL: D90730; BAA35675.1; -;
DR EMBL: D90731; BAA35681.1; -;
DR PIR: A03431; MMECF.
DR PIR: A25029; A25029.
DR PDB: ZOMP; 07-DEC-95.
DR PDB: LOFF; 07-FEB-95.
DR PDB: IGFN; 07-DEC-96.
DR PDB: IGFN; 07-DEC-96.
DR PDB: IGFO; 07-DEC-96.
DR PDB: IGFP; 07-DEC-96.
DR PDB: IGFO; 07-DEC-96.
DR PDB: IMPF; 07-FEB-95.
DR PDB: 1BT9; 13-JAN-99.
DR SWISS-2DPAGE; P02931; COLI.
DR ECO2DBASE; B036.0; 6TH EDITION.
DR EcoGene; EG10671; ompF.
DR InterPro; IPR003223; OMP_2.
DR InterPro; IPR001702; Porin_gram-ve.
DR Pfam; PF00267; Gram-ve_porins; 1.
DR PRINTS; PR00182; ECOLNEIPORIN.
DR PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
KW 3D-structure; Complete proteome.
FT SIGNAL 1 22
FT CHAIN 23 362 OUTER MEMBRANE PROTEIN F.
FT FT 88 Q -> E (IN REF. 5).
FT CONFLICT 139 139 E -> G (IN REF. 5).
FT CONFLICT 284 284 Q -> L (IN REF. 5).
FT STRAND 24 28
FT TURN 29 30
FT STRAND 31 45

```
FT TURN 50 51
FT STRAND 53 52
FT STRAND 58 59
FT STRAND 62 72
FT STRAND 77 88
FT TURN 95 100
FT STRAND 102 112
FT TURN 113 115
FT STRAND 116 124
FT TURN 126 127
FT HELIX 128 131
FT TURN 132 132
FT TURN 143 144
FT TURN 147 148
FT TURN 150 151
FT STRAND 154 163
FT HELIX 165 169
FT TURN 171 172
FT STRAND 173 180
FT STRAND 183 183
FT TURN 188 190
FT STRAND 192 192
FT STRAND 195 204
FT TURN 205 206
FT STRAND 207 217
FT HELIX 220 223
FT TURN 224 224
FT STRAND 227 227
FT STRAND 232 244
FT TURN 245 246
FT STRAND 247 257
FT TURN 261 264
FT STRAND 265 268
FT STRAND 269 272
FT STRAND 275 285
FT TURN 288 289
FT STRAND 291 305
FT TURN 306 308
FT STRAND 309 324
FT STRAND 329 338
FT STRAND 351 351
FT STRAND 353 361
SQ SEQUENCE 362 AA; 39333 MW; 3F0974D96DB65464 CRC64;

Query Match 8.9%; Score 85.5; DB 1; Length 362;
Best Local Similarity 25.2%; Pred. No. 1.1;
Matches 36; Conservative 22; Mismatches 60; Indels 25; Gaps 6;

QY 4 GVLGPYIGKSLTKLPAYWELSGFHQLTDQWA-----IHSYKYTEWSRFEKELRG 54
DQ 208 GIVGAY-GAADRITNLQEA--QPLNGKKAEQWATGLKYDANNIYLAANYGETRNPITN 264
QY 55 KYODGSGVEAFTKKEEYKDNRSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPD 114
DQ 265 KFTTSGFANKQDVLVAQYQDFEG-----LRPSIATYKSAKD--VEGIGDVD 312
QY 115 RM-WYSIGATYKFTPNLSVDVGF 136
DQ 313 LVNYFEVGATYTFNKNMSTYVDY 335

RESULT 12
HPUB_NEIMA
ID HPUB_NEIMA STANDARD; PRT; 810 AA.
AC Q9JMA2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemoglobin-haptoglobin utilization protein B precursor.
GN HPUB OR NMA0474.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OC NCBI_TaxID=65699;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RL meningitidis 22491."
RL Nature 404:502-506(2000).
CC -!- FUNCTION: ACTS AS A RECEPTOR FOR HEMOGLOBIN OR THE
CC HEMOGLOBIN/HAPTOGLOBIN COMPLEX AND IS REQUIRED FOR HEME UPTAKE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC -!- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL162753; CAB83769.1; -
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF00593; TonB_boxC; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
DR Outer membrane; Transport; TonB box; Signal; Receptor;
KW Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 810 HEMOGLOBIN-HAPTOGLOBIN UTILIZATION
FT PROTEIN B.
FT SITE 793 810 TONB C-TERMINAL BOX.
SQ SEQUENCE 810 AA; 90570 MW; D38DE1DCA3CA5A6E CRC64;

Query Match 8.8%; Score 84; DB 1; Length 810;
Best Local Similarity 24.2%; Pred. No. 4;
Matches 47; Conservative 17; Mismatches 74; Indels 56; Gaps 10;

QY 22 YWELSGFHQLTDQWAIHY-----SYKYTEWSRFEKELRGKYQD--GSGYEAFTK----- 67
DQ 378 YHSFRHQRNTAOWTADFKEKOLFSAKAWAAYGLGGKGDNANSYSYFAKLYDPKILA 437
QY 68 -----KKEYKD---NSRFAIGTYSLNDALTLRAGLAYDKAASK-----T 104
DQ 438 SNQAKITMLIENRSKYKFAWNNAPHLGG-----NDRFRLNAGIRYDKNSSAKDDPKYTT 493
QY 105 HLSASIP--DTPDR---MWYSIGATYKFTPNLSVDVGFHRLRGKK--HFVQTQNIKGLL 156
DQ 494 AIRGOIPHLSGERAHAGFSYCTGDFWRFTKHL-----HLLAKYSTGFRAPTSDETWLL 546
QY 157 LVEADYTTKATNL 170
DQ 547 FPHDPFVYKANKPNL 560

RESULT 13
ARHY_RAT
ID ARHY_RAT STANDARD; PRT; 362 AA.
AC Q02589;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADP-ribose-L-arginine
DE ADP-ribose-L-arginine
DE cleaving enzyme).
GN ADPRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


RA Whitehead S., Spratt B.G., Barrell B.G.;
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*
 RL meningitidis 22491";
 RN Nature 404:502-506(2000).
 [2]
 RP IDENTIFICATION OF THE METHIONINE SULFOXIDE REDUCTASE ACTIVITIES (MSRA
 RP AND MSRB).
 RC STRAIN=22491 / Serogroup A / Serotype 4A;
 RA Orly A., Boschi-Muller S., Warraud M., Sanglier-Cianferani S.,
 van Dorsselaar A., Brabant G.;
 RT "Characterization of the methionine sulfoxide reductase activities
 of PilB, a probable virulence factor from *Neisseria meningitidis*.";
 RL J. Biol. Chem. 277:10000-10006(2002).
 CC -1- FUNCTION: Has an important function as a repair enzyme for
 CC proteins that have been inactivated by oxidation (By similarity).
 CC Catalyzes the reversible oxidation-reduction of methionine
 CC sulfoxide in proteins to methionine.
 CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
 CC protein L-methionine S-oxide + reduced thioredoxin.
 CC -1- DOMAIN: Possesses 2 methionine sulfoxide reductase domains (A/MSrA
 CC and B/MSrB) and 1 N-terminal thioredoxin domain. The domain B
 CC exhibits a thioredoxin dependent methionine sulfoxide reductase
 CC activity; the Cys-495 is probably involved in the reduction of
 CC MetSO and in formation of the sulfenic acid derivative. The
 CC regeneration of Cys-495 is probably done via formation of a
 CC disulfide bond with Cys-440 followed by its reduction by
 CC thioredoxin.
 CC -1- MISCELLANEOUS: The domain msrB is stereospecific for the R isomer
 CC of the sulfoxide of MetSO whereas the domain msrA is
 CC stereospecific for the S isomer.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOREDOXIN
 CC FAMILY.
 CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE MSRA MET
 CC SULFOXIDE REDUCTASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE MSRB MET
 CC SULFOXIDE REDUCTASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: ALJ162752; CAB83597.1; -
 CC HSP: P54149; IPIVG.
 CC InterPro: IPR002579; DUF25.
 CC InterPro: IPR002569; PMSR.
 CC InterPro: IPR000063; ThioRed.
 CC Pfam: PF01641; DUF25; 1.
 CC Pfam: PF01625; PMSR; 1.
 CC ProDom: PD003489; PMSR; 1.
 CC ProDom: PD004057; DUF25; 1.
 CC TIGRfams: TIGR00357; DUF25; 1.
 CC TIGRfams: TIGR00401; msrA; 1.
 CC Oxidoreductase; Redox-active center; Electron transport;
 CC Multifunctional enzyme; Complete proteome.
 CC DOMAIN 38 171 THIOREDOXIN.
 CC DOMAIN 199 354 PEPTIDE METHIONINE SULFOXIDE REDUCTASE A.
 CC DOMAIN 383 506 PEPTIDE METHIONINE SULFOXIDE REDUCTASE B.
 CC ACT_SITE 207 207 BY SIMILARITY.
 CC FT DISULFID 68 71 REDOX-ACTIVE (BY SIMILARITY).
 CC FT DISULFID 440 495 PROBABLE.
 CC SQ SEQUENCE 522 AA; 58015 MW; F61E8EA7189F0667 CRC64;
 Query Match 8.6%; Score 82.5; DB 1; Length 522;
 Best Local Similarity 24.4%; Pred. No. 3.2;
 Matches 41; Conservative 22; Mismatches 56; Indels 49; Gaps 8;
 QY 11 GKGLPAYWLSGF-----HQLTDQWATHYSKYKWTENSREFK----- 50
 DB 370 GKGFDAATYKKPSDAELKRLTTEBOYOTNSATEYAFSHEYDHLFRPGIYDVVWSGEPL 429

QY 51 -ELRGKYDGSYGAEFTKKEEYKDNSRFAICTTYSNLNALTLAGLAYD-----KAAS 102
 DB 430 FSSADKYDSCGWPSPFRPIDAK-----SVTEHDDFSYNNRRTEVHSHAA 474
 QY 103 KTHLSASIPDTR----MWYSI-GATYKFTPNLSVD-VGFAHLRGKKK 144
 DB 475 DSHLGHVFPDGPDRKGLRYCINGASLKFTPLEQMDAAGYALKSKVK 522
 RESULT 16
 ARMY_MOUSE STANDARD; PRT: 362 AA.
 ID ARMY_MOUSE
 AC P54923;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADP-ribosylarginine hydrolase (EC 3.2.2.19) (ADP-ribose-L-arginine
 GN cleaving enzyme).
 GN ADPRH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93352593; PubMed=8349667;
 RA Takada T., Iida K., Moss J.;
 RT "Cloning and site-directed mutagenesis of human ADP-ribosylarginine
 RT hydrolase";
 RL J. Biol. Chem. 268:17837-17843(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Aoki K., Shoemaker M., Moss J.;
 RT "Genomic organization and promoter analysis of the mouse
 RT ADP-ribosylarginine hydrolase gene";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES REVERSE REACTION OF MONO-ADP-RIBOSYLATION.
 CC -1- CATALYTIC ACTIVITY: N(2)-(ADP-D-ribosyl)-L-arginine + H(2)O = L-
 CC arginine + ADP-ribose.
 CC -1- ENZYME REGULATION: SYNERGISTICALLY STIMULATED BY MAGNESIUM AND
 CC DITHIOERITOL (DTT) IN VITRO.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: LI3290; AAA37259.1; -
 CC DR EMBL: AF244347; AAF86223.1; -
 CC DR EMBL: BC003437; AAH03437.1; -
 CC DR MGD: MGI:1098234; Adprh.
 CC KW Hydrolase; Magnesium.
 CC SQ SEQUENCE 362 AA; 40068 MW; E06BCAEF69BA7C4E CRC64;
 Query Match 8.5%; Score 81.5; DB 1; Length 362;
 Best Local Similarity 27.2%; Pred. No. 2.6;
 Matches 37; Conservative 18; Mismatches 54; Indels 27; Gaps 8;
 QY 11 GKSLTKLP---AYWELSGFHQLTDQWATHYSKYKWTENSREFKRYQDGSGYEFTK 67
 DB 197 GKGLMEV-LPEAKKYITQSGY--FVKENLQHWYSYFEKEWKELELRC-ILDGNSAPVFPQ 252
 QY 68 KEEYKDNSRAICTTYS-----LNDALTLAGLAYDKAASKTHLSASIPD 114
 DB 253 PFGVKERDQFYDVSYSGWGGSGHGDAPMTAYDAL-LAAGDSWKELAIRAFFGGSDST 311

NMPC_ECOLI
 ID NMPC_ECOLI STANDARD: PRT: 365 AA.
 AC P21420; P77189;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Outer membrane porin protein nmPC precursor.
 GN NMPC OR PHMA OR B0553.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mutant CS384;
 RX MEDLINE=86304457; PubMed=3017988;
 RA Blasband A.J., Marcotte W.R. Jr., Schnaitman C.A.;
 RT "Structure of the lc and nmPC outer membrane porin protein genes of
 RL lambdaoid bacteriophage";
 RL J. Biol. Chem. 261:12723-12732(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick J.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
 RA Federspiel N., Hymen R., Kaiman S., Komp C., Kurdi O., Lcw H., Lin D.,
 RA Naphth A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-17 FROM N.A.
 RC STRAIN=JL5502;
 RX MEDLINE=94335635; PubMed=8057841;
 RA Coll J.L., Heyde M., Portallier R.;
 RT "Expression of the nmPC gene of Escherichia coli K-12 is modulated by
 external pH. Identification of cis-acting regulatory sequences
 RT involved in this regulation";
 RL Mol. Microbiol. 12:83-93(1994).
 RN [5]
 RP SEQUENCE OF 347-365 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=96196428; PubMed=8648624;
 RA Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
 RT "Holliday junction resolvases encoded by homologous rusa genes in
 RL Escherichia coli K-12 and phage 82";
 RL J. Mol. Biol. 257:561-573(1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- MISCELLANEOUS: IN WILD-TYPE STRAINS OF E. COLI K12, THE NMPC OPEN
 CC READING FRAME IS INTERRUPTED BY AN ISS INSERTION AND GENERATES A
 CC HYBRID OPEN READING FRAME THAT IS NOT EXPRESSED. HOWEVER, IN
 CC MUTANT STRAIN CS348, THE ISS ELEMENT HAS BEEN DELETED AND NMPC IS
 CC EXPRESSED.
 CC -1- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M13457; AAA23728.1; ALT_SEQ.
 DR EMBL; AE000160; AAC73654.1; ALT_SEQ.
 DR EMBL; U82598; AAB40749.1; ALT_SEQ.
 DR EMBL; Z35442; CAA84594.1;

DR EMBL; X92587; CAA63325.1; -;
 DR PIR; A25647; MMECNC.
 DR HSP; Q48473; IOSM.
 DR EcGene; EG10659; nmPC.
 DR InterPro; IPR003229; OMP_2.
 DR InterPro; IPR001702; Porin_gram-ve.
 DR Pfam; PF00267; Gram-ve_porins; 1.
 DR PRINTS; PR00182; ECOLNEIPORIN.
 DR PROSITE; PS00576; GRAM_NEG_PORIN; 1.
 KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
 FT SIGNAL 1
 FT CHAIN 24 365 OUTER MEMBRANE PORIN PROTEIN NMPC.
 FT CONFLICT 326 326 K -> N (IN REF. 2 AND 3).
 SQ SEQUENCE 365 AA; 40316 MW; 6E5128D4847FB4F8 CRC64;

 Query Match 8.4%; Score 80; DB 1; Length 365;
 Best Local Similarity 22.1%; Pred. No. 3.6;
 Matches 49; Conservative 29; Mismatches 70; Indels 74; Gaps 10;

 QY 21 AYWELSGFHQLTDD--QWAIHYSYK--YTEWSRFKELRGKYQDGSYGFAFTKKEEYKDSNR 76
 DB 155 ATYRNNDFFGLVDGLNFAAQYQGNDRSDFDNYTEGNG---DGFQFSA---TYEYEG--- 205
 QY 77 FAICTTYSLND-----ALTLRAGLAYDKAAKTHLSASIPDTR 115
 DB 206 FGIGATYAKSDRTDQVNAAGKLPVPEFASGKNAEVAAGLKID--ANNIYLATTTSETON 263
 QY 116 MW-----YSIGATYKFTPNLSVDVGFALHURGK----- 142
 DB 264 MTFADHFVANKAQNFAVAQYQDFGLRPSVAYLQSGKDLGVWGQDQLVKYVDVGATY 323
 QY 143 --KKHFVETONIKGLLLVEADYTTK---ATANLYGLNLNYRF 179
 DB 324 YFKKNMSTFVDYKINLLDKNDFTKALGVSTDDIVAGLVLYQF 365

 RESULT 20
 TRFE_HUMAN STANDARD; PRT: 698 AA.
 AC P02787; Q9UHV0; Q9NOB8; O43890;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serotransferrin precursor (transferrin) (Siderophilin) (Beta-1-metal
 DE binding globulin) (PRO1400).
 GN TF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS TF*B2; TF*CHI AND TF*DI.
 RX MEDLINE=84194084; PubMed=6585826;
 RA Yang F., Lum J.B., McGill J.R., Moore C.M., Naylor S.L.,
 RA van Bragt P.H., Baldwin W.D., Bowman B.H.;
 RT "Human transferrin: cDNA characterization and chromosomal
 RT localization";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:2752-2756(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88056305; PubMed=3678832;
 RA Schaeffer E., Lucero M.A., Jeitsch J.-M., Py M.-C., Levin M.J.,
 RA Chambon P., Cohen G.N., Zakim M.M.;
 RT "Complete structure of the human transferrin gene. Comparison with
 RT analogous chicken gene and human pseudogene";
 RL Gene 56:109-116(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92231399; PubMed=1809186;
 RA Hershsberger C.L., Larson J.L., Arnold B., Rosleck P.R. Jr.,
 RA Williams P., Dehoff B., Dunn P., O'Neal K.L., Riemen M.W.,
 RA Tice P.A.;

RT "A cloned gene for human transferrin.";
 RL Ann. N.Y. Acad. Sci. 646:140-154(1991).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT ATRANSFERRINEMIA PRO-477.
 RX MEDLINE=20563920; PubMed=11110675;
 RA Beutler E., Gelbart T., Lee P.L., Trevino R., Fernandez M.A.,
 RA Fairbanks V.F.;
 RT "Molecular characterization of a case of atransferrinemia.";
 RL Blood 96:4071-4074(2000).
 RN [5]
 RP SEQUENCE OF 99-698 FROM N.A.
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Bi J., Zhang Y., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 33 new genes deduced
 by analysis of cDNA clones from human fetal liver.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 422-698 FROM N.A.
 RX MEDLINE=84153910; PubMed=6322780;
 RA Uzan G., Frail M., Park I., Besmond C., Maessen G., Trepap J.S.,
 RA Zaklin M.M., Kahn A.;
 RT "Molecular cloning and sequence analysis of cDNA for human
 transferrin.";
 RL Biochem. Biophys. Res. Commun. 119:273-281(1984).
 RN [7]
 RP SEQUENCE OF 20-698.
 RX MEDLINE=83160878; PubMed=6833213;
 RA McGillivray R.T.A., Mendez E., Shewale J.G., Sinha S.K.,
 RA Lineback-Zins J., Brew K.;
 RT "The primary structure of human serum transferrin. The structures of
 seven cyanogen bromide fragments and the assembly of the complete
 structure.";
 RL J. Biol. Chem. 258:3543-3553(1983).
 RN [8]
 RP SEQUENCE OF 73-698 FROM N.A.
 RX MEDLINE=85216459; PubMed=3858812;
 RA Park I., Schaeffer E., Sildoll A., Baralle F.E., Cohen G.N.,
 RA Zaklin M.M.;
 RT "Organization of the human transferrin gene: direct evidence that it
 originated by gene duplication.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:3149-3153(1985).
 RN [9]
 RP SEQUENCE OF 1-14 FROM N.A.
 RX MEDLINE=87066744; PubMed=3786138;
 RA Lucero M.A., Schaeffer E., Cohen G.N., Zaklin M.M.;
 RT "The 5' region of the human transferrin gene: Structure and potential
 regulatory sites.";
 RL Nucleic Acids Res. 14:8692-8692(1986).
 RN [10]
 RP SEQUENCE OF 1-72 AND 291-300 FROM N.A.
 RX MEDLINE=87192006; PubMed=3106157;
 RA Adrian G.S., Korinek B.W., Bowman B.H., Yang F.;
 RT "The human transferrin gene: 5' region contains conserved sequences
 which match the control elements regulated by heavy metals,
 glucocorticoids and acute phase reaction.";
 RL Gene 49:167-175(1986).
 RN [11]
 RP SEQUENCE OF 45-72 FROM N.A.
 RX PubMed=10931525;
 RA de Arriba Zepa G.A., Saleh M.C., Fernandez P.M., Guillou F.,
 RA Espinosa de los Monteros A., de Vellis J., Zaklin M.M., Baron B.;
 RT "Alternative splicing prevents transferrin secretion during
 differentiation of a human oligodendrocyte cell line.";
 RL J. Neurosci. Res. 61:388-393(2000).
 RN [12]
 RP SEQUENCE OF 564-624 FROM N.A., AND VARIANT TP*C2.
 RC TISSUE=Brain;
 RX MEDLINE=97418135; PubMed=9272172;
 RA Namekata K., Oyama F., Imagawa M., Ihara Y.;
 RT "Human transferrin (Tf): a single mutation at codon 570 determines Tf
 C1 or Tf C2 variant.";
 RL Hum. Genet. 100:457-458(1997).

RN [13]
 RP SEQUENCE OF 564-624 FROM N.A.
 RA Tsuchida S., Ikemoto S., Kajii E.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [14]
 RP SEQUENCE OF 636-696 FROM N.A.
 RX MEDLINE=89386721; PubMed=2780570;
 RA Duguid J.R., Bohmont C.W., Liu N.G., Tourtellotte W.W.;
 RT "Changes in brain gene expression shared by scrapie and Alzheimer
 disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7260-7264(1989).
 RN [15]
 RP SEQUENCE OF 263-266; 454-458; 531-538 AND 589-595.
 RC TISSUE=Heart;
 RX MEDLINE=96007936; PubMed=7498159;
 RA Kovalyov L.I., Shishkin S.S., Efilochkin A.S., Kovalyova M.A.,
 RA Ershova E.S., Egorov T.A., Musalyanov A.K.;
 RT "The major protein expression profile and two-dimensional protein
 database of human heart.";
 RL Electrophoresis 16:1160-1169(1995).
 RN [16]
 RP DISULFIDE BONDS
 RX MEDLINE=82222166; PubMed=6953407;
 RA McGillivray R.T.A., Mendez E., Sinha S.K., Sutton M.R.,
 RA Lineback-Zins J., Brew K.;
 RT "The complete amino acid sequence of human serum transferrin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2504-2508(1982).
 RN [17]
 RP MUTAGENESIS.
 RX MEDLINE=92031536; PubMed=1932003;
 RA Woodworth R.C., Mason A.B., Funk W.D., McGillivray R.T.A.;
 RT "Expression and initial characterization of five site-directed
 mutants of the N-terminal half-molecule of human transferrin.";
 RL Biochemistry 30:10824-10829(1991).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 22-350.
 RX MEDLINE=98272665; PubMed=9609685;
 RA Macgillivray R.T.A., Moore S.A., Chen J., Anderson B.F., Baker H.,
 RA Luo Y., Bewley M.C., Smith C.A., Murphy M.E.P., Wang Y., Mason A.B.,
 RA Woodworth R.C., Brayer G.D., Baker E.N.;
 RT "Two high-resolution crystal structures of the recombinant N-lobe of
 human transferrin reveal a structural change implicated in iron
 release.";
 RL Biochemistry 37:7919-7928(1998).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 22-350.
 RX MEDLINE=98434369; PubMed=9760232;
 RA Jeffrey P.D., Bewley M.C., Macgillivray R.T.A., Mason A.B.,
 RA Woodworth R.C., Baker E.N.;
 RT "Ligand-induced conformational change in transferrins: crystal
 structure of the open form of the N-terminal half-molecule of human
 transferrin.";
 RL Biochemistry 37:13978-13986(1998).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 23-352.
 RX MEDLINE=99155227; PubMed=10029548;
 RA Bewley M.C., Tam 8.M., Grewal J., He S., Shewry S., Murphy M.E.P.,
 RA Mason A.B., Woodworth R.C., Baker E.N., Macgillivray R.T.A.;
 RT "X-ray crystallography and mass spectroscopy reveal that the N-lobe
 of human transferrin expressed in *Pichia pastoris* is folded correctly
 but is glycosylated on serine-32.";
 RL Biochemistry 38:2535-2541(1999).
 RN [21]
 RP VARIANT SER-142.
 RX MEDLINE=98019079; PubMed=9358047;
 RA Evans P., Kemp J.;
 RT "Exon/intron structure of the human transferrin receptor gene.";
 RL Gene 199:123-131(1997).
 RN [22]
 RP VARIANT GLU-646.
 RX PubMed=9803271;
 RA Pang H., Koda Y., Soejima M., Kimura H.;
 RT "Identification of a mutation (Al879G) of transferrin from cDNA

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss;
 RX MEDLINE=93096477; PubMed=1334250;
 RA Reynolds A.B., Herbert L., Cleveland J.L., Berg S.T., Gant J.R.;
 RT "p120, a novel substrate of protein tyrosine kinase receptors and of
 RT p60v-src, is related to cadherin-binding factors beta-catenin,
 RT plakoglobin and armadillo";
 RL Oncogene 7:2439-2445(1992).
 RN [2]
 RP REVISIONS.
 RA Reynolds A.B.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EFFICIENT TYROSINE KINASE SUBSTRATE IMPLICATED BOTH IN
 CC CELL TRANSFORMATION BY SRC AND IN LIGAND-INDUCED RECEPTOR (BY
 CC SIGNALING THROUGH THE EGF, PDGF, CSF-1 AND EHRB2 RECEPTORS (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL ADHESION
 CC MOLECULE E-CADHERIN. THE ASSOCIATION OF CATEININS TO CADHERINS
 CC PRODUCES A COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK,
 CC AND WHICH SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-
 CC ADHESION PROPERTIES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR IN PARTICULAR CELLS
 CC (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED BY PROTEIN-TYROSINE KINASES.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: 217804; CAA79078.1; -
 DR PIR: S28498; S28498.
 DR MGD: MGI:105100; Catns
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 5.
 DR SMART: SM00185; ARM; 4.
 DR PROSITE: PS50176; ARM_REPEAT; 3.
 DR Cytoskeleton; Structural protein; phosphorylation; repeat;
 KW Cell adhesion; Coiled coil; Nuclear protein.
 FT DOMAIN 10 46 COILED COIL (POTENTIAL).
 FT REPEAT 358 395 ARM 1.
 FT REPEAT 398 437 ARM 2.
 FT REPEAT 441 475 ARM 3.
 FT REPEAT 476 516 ARM 4.
 FT REPEAT 534 573 ARM 5.
 FT REPEAT 583 624 ARM 6.
 FT REPEAT 647 687 ARM 7.
 FT REPEAT 694 733 ARM 8.
 FT REPEAT 734 774 ARM 9.
 FT REPEAT 775 820 ARM 10.
 FT DOMAIN 622 628 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 911 AA; 101731 MW; EE18C623948DD01 CRC64;
 Query Match 8.3%; Score 79; DB 1; Length 911;
 Best Local Similarity 25.2%; Pred. No. 13;
 Matches 29; Conservative 15; Mismatches 43; Indels 28; Gaps 6;
 QY 3 NGVLGPIYIGKSLTLKLPAYNELSGFHQLTDQWAIHY-----SYKYTEWSRFEKLRGK 55
 Db 187 NGPGPYVGQAG-TATLP-----RNFHYPPDGYGRHYEDGYGGSDNVSURVTRIER 240
 QY 56 YQDG-SGYEAFKKEEYKDNRSFAIGTF-----YSINDALTLRAGLADY 98
 Db 241 YRPMSEGRYRPSQDVGYPQVRVGGSSVDLHRFPPEYGLEDD---QRSGVD 292

RESULT 24
 PAGC_SALTY STANDARD; PRT; 185 AA.
 AC P23988;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Virulence membrane protein pagC precursor.
 GN PAGC OR STM1246.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14028;
 RX MEDLINE=91100323; PubMed=1846140;
 RA Puikkinen W.S., Miller S.I.;
 RT "A Salmonella typhimurium virulence protein is similar to a Yersinia
 RT enterocolitica invasion protein and a bacteriophage lambda outer
 RT membrane protein";
 RL J. Bacteriol. 173:86-93(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2".
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: ESSENTIAL FOR FULL VIRULENCE AND SURVIVAL WITHIN
 CC MACROPHAGES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- SIMILARITY: BELONGS TO THE AIL/OMPX/PAGC/LOM FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M55546; AAA27179.1; -
 DR EMBL: AE008754; AAL20174.1; -
 DR PIR: A39185; A39185.
 DR HSP: P36546; IQJ8.
 DR StyGene: SG10676; pagC.
 DR InterPro: IPR000758; Enterovir_OMP.
 DR PROSITE: PS00694; ENT_VIR_OMP_1; 1.
 DR PROSITE: PS00695; ENT_VIR_OMP_2; 1.
 FT Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 185 VIRULENCE MEMBRANE PROTEIN PAGC.
 FT CONFLICT 33 33 A -> ARYA (IN REF. 1).
 FT CONFLICT 185 AA; 20183 MW; 10B4E785032ECB9E CRC64;
 SQ SEQUENCE 185 AA; 20183 MW; 10B4E785032ECB9E CRC64;
 Query Match 8.2%; Score 78.5; DB 1; Length 185;
 Best Local Similarity 23.9%; Pred. No. 2.1;
 Matches 45; Conservative 20; Mismatches 56; Indels 67; Gaps 10;
 QY 33 DQWAIHVSYYKTEWSRFEKLRG-----KYQDGSYEAFTKKEEYKDNRSFA--2----- 78
 Db 24 DTNAFSGVGAQSGKQVDFKNGIRGVNVKRYEDDSPV-SFISSLSLYGDRQASGVPEGI 82
 QY 79 -----IGTYSINDALTL--RAGLAYDAASKTHLSASIPDTD----- 114
 Db 83 HYHDKFEVYKGLVMGPAYRLSDNFSLYALAGVGTVKATKEH---STQDGDSTFSNKISS 139


```

Oy 115 -RMWYSIGATYKTP--NLSVDVGFALRGKKHFEVETQIKGLLVEADYTTKATANLV 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 RKTGFAGAGVQNNPLENVVDVGY-----EGSNIS-----STK--INGF 177
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 172 GLNLNTRF 179
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 NVGVGYRF 185

RESULT 25
Y86A_HAEIN STANDARD: PRT: 345 AA.
AC 086241;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical tonB-dependent receptor H1466.1.
GN H1466.1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Wellman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
RN [2]
RP REVISIONS.
RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.,
RA Peterson J., Hickey E., Dodson R., Gwinn M.J.
RA Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U32825; AAC23131.1;
CC TIGR; H1466.1;
CC InterPro: IPR000531; TonB_boxC.
CC Pfam; PF00593; TonB_boxC; 1.
CC PROSITE; PS01156; TONB_DEPENDENT_REC_2; FALSE_NEG.
KW Hypothetical protein; Receptor; TonB box; Complete proteome.
FT SITE 325 345 TONB C-TERMINAL BOX.
SQ SEQUENCE 345 AA; 38744 MW; EAF83C55FCC9CB2A CRC64;

Query Match 8.2%; Score 78.5; DB 1; Length 345;
Best Local Similarity 25.7%; Pred. No. 4.6;
Matches 35; Conservative 17; Mismatches 53; Indels 31; Gaps 7;

Oy 67 KKEE-----YKDNRSFAIG-----TTVSLNDALTLRAGLAYD-----KAASKTHLSASI 110
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 KKNENNNAQFNHLKTLGLYIQNTTY-FTDNFIITGLRYEYFDQVVGSTLKNIRSGYL 90
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 111 PSTD-RMWYSIGATYKTPNLSVDVGFALRGKKHFEVETQIKGLLVEADY----- 163
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 91 AOKDGLLYOLGSGVYKFTPNIA--TFNHAESFRPNQNRLLIINGELPABQGSFETGLK 148
Oy 164 -----TKATANLYGLN 174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 YENAYLNATVALENNIN 164
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 26
FPTA_PSEAF STANDARD: PRT: 720 AA.
AC P42512;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fe(III)-pyochelin receptor precursor.
GN FPTA OR PA4221.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-51.
RC STRAIN=PAO / Ia602;
RX MEDLINE=94117363; PubMed=8288523;
RA Ankenbauer R.G., Quan H.N.;
RT "Fpca, the Fe(III)-pyochelin receptor of Pseudomonas aeruginosa: a
RT phenolate siderophore receptor homologous to hydroxamate siderophore
RT receptors.";
RL J. Bacteriol. 176:307-319(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Garber R.L., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: HIGH-AFFINITY OUTER MEMBRANE RECEPTOR REQUIRED FOR THE
CC TRANSPORT OF FE(III)-PYOCHELIN.
CC -1- SUBCELLULAR LOCATION: Outer membrane.
CC -1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U03161; AAC43213.1;
CC EMBL; AE004839; AAG07609.1;
CC InterPro: IPR000531; TonB_boxC.
CC Pfam; PF00593; TonB_boxC; 1.
CC PROSITE; PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
CC PROSITE; PS01156; TONB_DEPENDENT_REC_2; 1.
KW Outer membrane; Iron transport; Transport; Signal; Receptor;
KW Complete proteome.
FT SIGNAL 1 38
FT CHAIN 39 720 FE(III)-PYOCHELIN RECEPTOR.
FT SITE 703 720 TONB C-TERMINAL BOX.
SQ SEQUENCE 720 AA; 79992 MW; DA796313116E0CC2 CRC64;

Query Match 8.2%; Score 78.5; DB 1; Length 720;
Best Local Similarity 24.0%; Pred. No. 11;
Matches 35; Conservative 17; Mismatches 47; Indels 47; Gaps 8;

Oy 23 WELSGPHQLTDQWAIHVSYKYTEWSRFKELRGKYQDGSYGAEFTKKEYSRFAIGTT 82

```

```

Db 577 FELEGTGYLTYPWLSAGYTYTSTYLYLD--SQNDSGFRYSTFTPRH----- 621
Qy 83 YSLNDALTLRAGLAYDKAASHTLSASIPDTRMWSYIGATYKTPNLSVDV----- 134
Db 622 -----LRLWSNYD-----LPWQDRRW-SVGGGLQAQSDYSVDYRGVSMRQG 662
Qy 135 GFA--HLR-GRK--KHFEVETQIRGL 155
Db 663 GYALVNMRLGYKIDEHHTAAVNVNVL 688

RESULT 27
NX2A_RAT
ID NX2A_RAT STANDARD; PRT; 1715 AA.
AC Q63374; Q63375;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurexin 2-alpha precursor (Neurexin II-alpha).
OS NRXN2.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A., VARIANT LEU-434, AND ALTERNATIVE SPLICING.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92320296; PubMed=1621094;
RA Ushkaryov Y.A., Petrenko A.G., Geppert M., Suedhof T.C.;
RT "Neurexins: Synaptic cell surface proteins related to the
RL alpha-latrotoxin receptor and laminin.";
RN Science 257:50-56(1992).
[2]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=95209856; PubMed=7695896;
RA Ullrich B., Ushkaryov Y.A., Suedhof T.C.;
RT "Cartography of neurexins: more than 1000 isoforms generated by
RL alternative splicing and expressed in distinct subsets of neurons.";
RN Neuron 14:497-507(1995).
[3]
INTERACTION WITH NEUREXOPHILIN 1.
RX MEDLINE=99074239; PubMed=9856994;
RA Massier M., Hammer R.E., Suedhof T.C.;
RT "Neurexophilin binding to alpha-neurexins. A single LNS domain
RL functions as an independently folding ligand-binding unit.";
RN J. Biol. Chem. 273:34716-34723(1998).
[4]
INTERACTION WITH ALPHA-DYSTROGLYCAN.
RX MEDLINE=21363578; PubMed=11470830;
RA Sugita S., Saito F., Tang J., Satz J., Campbell K., Suedhof T.C.;
RT "A stoichiometric complex of neurexins and dystroglycan in brain.";
RL J. Cell Biol. 154:435-445(2001).
CC -!- FUNCTION: NEURONAL CELL SURFACE PROTEIN THAT MAY BE INVOLVED IN
CC CELL RECOGNITION AND CELL ADHESION. MAY MEDIATE INTRACELLULAR
CC SIGNALING.
CC -!- SUBUNIT: LAMININ G-LIKE DOMAIN 2 BINDS TO NEUREXOPHILIN 1.
CC ISOFORMS ALPHA 2C BIND TO ALPHA-DYSTROGLYCAN.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 216 ISOFORMS; MAY BE PRODUCED BY
CC SPLICED EXTRACELLULAR DOMAINS AT SITES 1 TO 5, EACH CONSISTING OF
CC MODULAR SEQUENCES (A-C) THAT SEEM TO BE USED INDEPENDENTLY.
CC ADDITIONAL ISOFORMS MAY DERIVE FROM A MINOR CYTOPLASMIC SPLICE
CC SITE 6. BETA ISOFORMS (AC Q63376) SHARE THE COMBINATION OF
CC ALTERNATIVE SPLICED DOMAINS AT SITES 4, 5 AND 6. SHOWN IS ISOFORM
CC ALPHA IA2A34A3A. ALPHA AND BETA ISOFORMS DIFFER IN THEIR N-
CC TERMINUS BY USE OF ALTERNATIVE PROMOTERS IN THE NRXN2 GENE.
CC -!- TISSUE SPECIFICITY: BRAIN (NEURONAL SYNAPSE).
CC -!- SIMILARITY: CONTAINS 6 LAMININ G-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE NEUREXIN FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
-----
EMBL; M63376; AAA11706.1; -.
DR EMBL; M63376; AAA11707.1; -.
DR HSP; Q63373; IC4R.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00008; EGF_3.
DR Pfam; PF00054; laminin_G; 5.
DR SMART; SM00181; EGF_3.
DR SMART; SM00282; LamG; 6.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00025; LAM_G_DOMAIN; 6.
KW Signal; Transmembrane; Repeat; Cell adhesion; Glycoprotein;
KW Alternative splicing; Polymorphism.
FT SIGNAL 1 29
FT CHAIN 30 1715
FT DOMAIN 30 1639
FT TRANSMEM 1640 1660
FT DOMAIN 1661 1715
FT DOMAIN 30 206
FT DOMAIN 202 242
FT DOMAIN 289 486
FT DOMAIN 493 686
FT DOMAIN 690 727
FT DOMAIN 732 907
FT DOMAIN 921 1096
FT DOMAIN 1099 1136
FT DOMAIN 1140 1348
FT DOMAIN 1371 1374
FT DOMAIN 1377 1380
FT DOMAIN 1447 1450
FT DOMAIN 1647 1650
FT CARBOHYD 60 60
FT CARBOHYD 295 295
FT CARBOHYD 338 338
FT CARBOHYD 844 844
FT CARBOHYD 1239 1239
FT VARSPPLIC 250 259
FT VARSPPLIC 250 283
FT VARSPPLIC 385 399
FT VARSPPLIC 393 399
FT VARSPPLIC 797 806
FT VARSPPLIC 807 809
FT VARSPPLIC 1256 1285
FT VARSPPLIC 1421 1614
FT VARSPPLIC 1667 1715
FT VARIANT 434 434 N->L.
FT SEQUENCE 1715 AA; 185282 MW; 59FBF18661F3DB15 CRC64;
Query Match 8.2%; Score 78.5; DB 1; Length 1715;
Best Local Similarity 20.4%; Pred. No. 32;
Matches 37; Conservative 25; Mismatches 60; Indels 59; Gaps 7;
Qy 54 GKYDGSQCYEAFKKEE---YKNSRFAICTTY-----SLNDALTL-----RAGLAYDK 99
Db 271 GRGAGVNHQTKKEEFVATFKGNESFCYDLNSHPNQSSTDEITLAFRTLQRGLMLHT 330
Qy 100 AAKSTHLSASIPD-----TDRMWSYIGATYKTPNLSVDV 134
Db 331 GKSANYNLSLKSGAVWLIINLGSQAEALVEPVNGRFNDNAHDV-----RVTRNLQHA 386

```


Db 217 GNPFVWKDTFLKNNRME---KHLRRKKGH 245

RESULT 30

PORT_BPPA2

ID_ PORT_BPPA2 STANDARD; PRT; 365 AA.

AC P07238;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Outer membrane porin protein LC precursor.

GN LC.

OS Bacteriophage PA-2.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;

OC unclassified Siphoviridae.

OX NCBI_TaxID=10738;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86304457; PubMed=3017988;

RA Blasband A.J., Marcotte W.R. Jr., Schnaitman C.A.;

RT "Structure of the lc and mmpC outer membrane porin protein genes of lambdaoid bacteriophage."

RL J. Biol. Chem. 261:12723-12732(1986).

CC -!- FUNCTION: PORINS ARE MAJOR PROTEINS FOUND IN THE OUTER MEMBRANES OF GRAM-NEGATIVE BACTERIA WHERE THEY FORM CHANNELS FOR THE NONSPECIFIC PERMEATION OF SMALL SOLUTES (MOLECULES WITH MW LOWER THAN 4000-6000 DALTONS).

CC -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane (By similarity).

CC -!- MISCELLANEOUS: A PORIN GENE CAN ALSO BE FOUND IN THE GENOMES OF CERTAIN LAMBDOID BACTERIOPHAGE, AND ITS PROTEIN IS EXPRESSED IN THE LYSOGENIC STATE. IN E.COLI THE EXPRESSION OF OMPC AND OMPF PROTEINS IS THEN REDUCED SUBSTANTIALLY.

CC -!- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; J02580; AAA32301.1; -

DR PIR; D25647; MMBPP2.

DR HSSP; O52503; IIIV.

DR InterPro; IPR003229; OMP_2.

DR InterPro; IPR001702; Porin-gram-ve.

DR Pfam; PF00267; gram-ve-porins; 1.

DR PRINTS; PR00182; ECOLNEIPORIN.

DR PROSITE; PS00576; GRAM_NEG_PORIN; 1.

KW Outer membrane; Transmembrane; Porin; Signal.

FT SIGNAL 1 23

FT CHAIN 24 365 OUTER MEMBRANE PORIN PROTEIN LC.

FT CONFLICT 99 99 H -> R (IN REF. 1; AA SEQUENCE).

SQ SEQUENCE 365 AA; 40290 MW; 0FBC0531FB9C0205 CRC64;

Query Match 8.0%; Score 77; DB 1; Length 365;

Best Local Similarity 20.2%; Pred. No. 6.8;

Matches 48; Conservative 27; Mismatches 57; Indels 106; Gaps 10;

Qy 21 AYWELSGPHQITD--QWAIHSYK--YTEMSEKRELKGYQDGSYEAFTEKEEYKNSR 76

Db 155 ATYRNDFEGLVDGLNFAAQYQGGKNDKSDFDNYTEGNG---DGGFSA---TYEVEG--- 205

Qy 77 FAIGTYSIND-----ALTRAGLAYDKAAKSTHLSASIPDTR 115

Db 206 FGIGATYAKSDRTQVWAGKVLPEVFASGKNAEVAAGLYD--ANNIYLATYSETQN 263

Qy 116 M-----WYSIGATY 124

Db 264 MTFVADHFVANKQNEAQAQYQDFGLRPSVAYLQSKGLGVWGDQDLVKYVDVGATY 323

QY 125 KFTPNLSVDVGFHRLGRKKKHEVETQNIKGLLLVADYTK---ATANLYGILNLYRF 179

Db 324 YFNKNWSIFVDY-----KINLLDNDFTKALGVSTDDDIVAVGLVQF 365

RESULT 31

FECA_ECOLI

ID FECA_ECOLI STANDARD; PRT; 774 AA.

AC P13036;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Iron(III) dicitrate transport protein feaA precursor.

GN FECA OR B4291.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-40.

RC STRAIN=B.

RX MEDLINE=88227855; PubMed=2836368;

RA Pressier U., Staudenmaier H., Zimmermann L., Braun V.;

RT "Genetics of the iron dicitrate transport system of Escherichia coli."

RL J. Bacteriol. 170:2716-2724(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=95334362; PubMed=7610040;

RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L., Blattner F.R.;

RT "Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes."

RL Nucleic Acids Res. 23:2105-2119(1995).

RN [3]

RP SEQUENCE OF 1-428 FROM N.A.

RX MEDLINE=91072220; PubMed=2254251;

RA Hove B., Staudenmaier H., Braun V.;

RT "Novel two-component transmembrane transcription control: regulation of iron dicitrate transport in Escherichia coli K-12."

RL J. Bacteriol. 172:6749-6758(1990).

RN [4]

RP SEQUENCE OF 731-774 FROM N.A.

RC STRAIN=K12;

RX MEDLINE=89213950; PubMed=2651410;

RA Staudenmaier H., van Hove B., Yaraghi Z., Braun V.;

RT "Nucleotide sequences of the fecBCE genes and locations of the proteins suggest a periplasmic-binding-protein-dependent transport mechanism for iron(III) dicitrate in Escherichia coli."

RL J. Bacteriol. 171:2626-2633(1989).

CC -!- FUNCTION: FECA IS THE OUTER MEMBRANE RECEPTOR PROTEIN IN THE IRON(III) DICITRATE TRANSPORT SYSTEM.

CC -!- SUBCELLULAR LOCATION: Outer membrane.

CC -!- INDUCTION: FOR INDUCTION THE TONB AND THE EXB PROTEIN HAVE TO BE ACTIVE. REGULATION BY THE IRON LEVEL MEDIATED BY THE FUR PROTEIN AND INDUCTION BY CITRATE PLUS IRON SUGGEST THAT THE IRON(III) DICITRATE COMPLEX MUST ENTER THE PERIPLASM WHERE IT BINDS TO A TRANSMEMBRANE PROTEIN, WHICH REGULATES DIRECTLY OR VIA A FURTHER INDUCTOR, TRANSCRIPTION OF THE FEC GENES.

CC -!- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M20581; AAA23760.1; -

DR EMBL; U14003; AAA97187.1; -

```
DR EMBL: A5000499; AAC77247.1; -.
DR EMBL: M63115; AAA23768.1; -.
DR EMBL: M26397; AAA23761.1; -.
DR PIR: JY0022; ORECEA.
DR EcoGene: EGI0286; fecA.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF00593; TonB_boxC; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
DR PROSITE; PS01156; TONB_DEPENDENT_REC_2; 1.
KW Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor;
KW Complete proteome.
FT SIGNAL 1 33
FT CHAIN 34 774
FT FT
FT SITE 56 63
FT SITE 757 774
FT CONFLICT 16 16
FT CONFLICT 190 190
FT CONFLICT 300 300
FT CONFLICT 357 357
FT CONFLICT 444 444
FT CONFLICT 749 749
FT CONFLICT 774 774
FT SEQUENCE 774 AA; 85321 MW; 16B5B510276C3B09 CRC64;
Query Match 8.0%; Score 76.5; DB 1; Length 774;
Best Local Similarity 20.7%; Pred. No. 19;
Matches 51; Conservative 32; Mismatches 68; Indels 95; Gaps 12;
QY 1 QHNGVLGPYIGKSLTLKLPAYWELSGFHOITDQWAIHYKKTYS----- 47
DB 508 QNNAITGTH-EEVSYNAPLPA---LNVLYHLTDSWNL---YANTEGSFGTVQVYSQIGKAV 560
QY 48 -----RFKELRGKYQDGS-----GYEAFTKKEEYKDN----- 74
DB 561 OSGNVEPEKARTWELGTRYDDGALTAEMGLFLNFNQYDSNQTNDVTARGKTRHTGLE 620
QY 75 --SREAITGTYSLNDALTLAGLAY-----DKAASKTHLSASIPDPDRMWSIGATYK- 125
DB 621 TQARYDLGLTPTLTDNVSIVAYVNAIREKRGDTYGNL---VPFSPKHGTLGVDPYK 677
QY 126 --FTPNLSVDVGFALHRLGKKHFEVETQNIKGLLLVEADYTTKATAN-----LYGL 173
DB 678 GNMFTFNLSDF-----QSSQFADNAN-----TVKESADGSTGRIPGFMLWGA 719
QY 174 NLNRYF 179
DB 720 RVAYDF 725
RESULT 32
PORI_RHOBL
ID PORI_RHOBL STANDARD; PRT; 289 AA.
AC P39767;
DT 01-FEB-1995 (Rel. 31, Created).
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Porin.
GN OPMA.
OS Rhodopseudomonas blautica.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1075;
RN [1]
RX MEDLINE-94191532; PubMed-8142898;
RA Kreusch A., Neubueser A., Schiltz E., Weckesser J., Schulz G.E.;
RT "Structure of the membrane channel porin from Rhodopseudomonas
RL blautica at 2.0-A resolution.";
RN Protein Sci. 3:58-63(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.96 ANGSTROMS).
RX MEDLINE-95055730; PubMed-7525973;
RA Kreusch A., Schulz G.E.;
```

```
RT "Refined structure of the porin from Rhodopseudomonas blautica.
RT Comparison with the porin from Rhodobacter capsulatus.";
RL J. Mol. Biol. 243:891-905(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF MUTANTS.
RX MEDLINE-98348035; PubMed-9684893;
RA Schmid B., Maveyraud L., Kromer M., Schulz G.E.;
RT "Porin mutants with new channel properties.";
RL Protein Sci. 7:1603-1611(1998).
CC -!- FUNCTION: FORMS CHANNELS THAT ALLOW THE PASSIVE DIFFUSION OF SMALL
CC HYDROPHILIC SOLUTES UP TO AN EXCLUSION LIMIT OF ABOUT 0.6 kDa.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
DR PIR: S38806; S38806.
DR PDB: 1PRN; 24-OCT-94.
DR PDB: 2PRN; 13-JAN-99.
DR PDB: 3PRN; 12-JAN-99.
DR PDB: 5PRN; 12-AUG-98.
DR PDB: 6PRN; 12-AUG-98.
DR PDB: 7PRN; 12-AUG-98.
DR PDB: 8PRN; 12-AUG-98.
DR PDB: 1BH3; 12-AUG-98.
DR InterPro: IPR001702; Porin_gram-ve.
DR Pfam: PF00267; Gram-ve_porins; 1.
KW Outer membrane; Transmembrane; Porin; 3D-structure.
SQ SEQUENCE 289 AA; 30597 MW; 08252D9803A1044C CRC64;
Query Match 7.9%; Score 76; DB 1; Length 289;
Best Local Similarity 32.2%; Pred. No. 63;
Matches 29; Conservative 7; Mismatches 28; Indels 26; Gaps 5;
QY 73 DNSRFAIGTYSLNDALTLAGLA-YDKAAS-----KTHLSASIPDTRM 116
DB 190 DNDIAFGAAYKFNDACTV--GLNWDNGLSTAGDQVTLGNYAFGATTVRAYVSDIDRA 247
QY 117 W-----YSIGATYKTPNL-----SVDVGFH 138
DB 248 GADTAYGIGADYQFAEGVKVSGVQSGFAN 277
RESULT 33
ARHY_HUMAN
ID ARHY_HUMAN STANDARD; PRT; 357 AA.
AC P54922;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADP-ribosylarginine hydrolase (EC 3.2.2.19) (ADP-ribose-L-arginine
DE cleaving enzyme).
GN ADPRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE-93352593; PubMed-8349667;
RA Takada T., Iida K., Moss J.;
RT "Cloning and site-directed mutagenesis of human ADP-ribosylarginine
RT hydrolase.";
RL J. Biol. Chem. 268:17837-17843(1993).
CC -!- FUNCTION: IT CATALYZES REVERSE REACTION OF MONO-ADP-RIBOSYLATION.
CC -!- CATALYTIC ACTIVITY: N(2)-(ADP-D-ribosyl)-L-arginine + H(2)O -> L-
CC arginine + ADP-ribose.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```



```

Db 292 LRPSLAYLQSKGK-NLGRGYDDEDLTKYVDVGATYFKNKMSYVDY----- 337
Qy 150 QNKGKLLLEADYTTKA---TANLYGLNLNRYF 179
Db 338 ---KINLLDNOFTRDAGINTDNIVALLGLVYQF 367

RESULT 35
OMPC_SALTY
ID OMPC_SALTY STANDARD; PRT; 378 AA.
AC Q52503; P09878;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Outer membrane protein C precursor (Porin ompC).
GN OMPC OR STM2267 OR STY2493.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=ATCC 14028;
RX MEDLINE=99457705; PubMed=10528398;
RA Negm R.S., Pistole T.G.;
RT "The porin OmpC of Salmonella typhimurium mediates adherence to
macrophages";
RL Can. J. Microbiol. 45:658-669(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Aili J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi;
RX MEDLINE=90060831; PubMed=2684785;
RA Puente J.L., Alvarez-Scherer V., Gosset G., Calva E.;
RT "Comparative analysis of the Salmonella typhi and Escherichia coli
ompC genes.";
RL Gene 83:197-206(1989).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2;
RX MEDLINE=88319959; PubMed=3412902;
RA Venegas A., Gomez I., Zaror I., Yudelevich A.;
RT "The nucleotide sequence of the Salmonella typhi ompC porin gene.";
RL Nucleic Acids Res. 16:7721-7721(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jageis K.,
Krogh A., Larsen T.S., Leachner S., Moule S., O'Gaora P., Parry C.,
Quail M., Rutherford K., Skellerns M., Skellern J., Stevens K.,
Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [6]

```

```

RP 3D-STRUCTURE MODELING, AND CRYSTALLIZATION.
RC SPECIES=S.typhi;
RX MEDLINE=99332091; PubMed=10405180;
RA Arockiasamy A., Krishnaswamy S.;
RT "Crystallization of the immunodominant outer membrane protein OmpC;
the first protein crystals from Salmonella typhi, a human pathogen.";
RL FEBS Lett. 453:380-382(1999).
CC -!- FUNCTION: FORMS PASSIVE DIFFUSION PORES WHICH ALLOW SMALL
MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE OUTER MEMBRANE.
CC -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -!- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF039309; AAB96675.1; -
CC EMBL; AE008801; AAL21169.1; -
CC EMBL; M31424; AAA27169.1; -
CC EMBL; X07835; CAA30688.1; -
CC EMBL; AL627274; CAD07499.1; -
CC PIR; JQ0119; MMEBPC.
CC PIR; S01248; S01248.
CC PDB; 1JIV; 14-OCT-98.
CC StyGene; SG10675; ompC.
CC InterPro; IPR003229; OMP_2.
CC InterPro; IPR001702; Porin_gram-ve.
CC Pfam; PF00267; Gram-ve_porins; 1.
CC PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
KW 3D-structure; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 378 OUTER MEMBRANE PROTEIN C.
FT CONFLICT 42 42 H -> N (IN REF. 1).
FT CONFLICT 223 223 N -> D (IN REF. 1).
FT CONFLICT 253 253 L -> V (IN REF. 1).
FT CONFLICT 258 258 S -> F (IN REF. 1).
FT CONFLICT 340 340 K -> T (IN REF. 1).
FT CONFLICT 356 357 ND -> YE (IN REF. 1).
FT CONFLICT 362 362 MISSING (IN REF. 4).
SQ SEQUENCE 378 AA; 41239 MW; 58479E8685B43ED1 CRC64;

Query Match 7.9%; Score 75.5; DB 1; Length 378;
Best Local Similarity 27.2%; Pred. No. 9.7;
Matches 41; Conservative 17; Mismatches 60; Indels 33; Gaps 8;

Qy 36 AIHYSYKYTEWSRFKELRGKYQDGSQGYEATFKKEEYKONSRAFGITTSINDALTLRAGL 95
Db 254 AAQYQSYTNA-TRFGTSNG--SNPSTSYGFANKAQ-----NFEVVAQYQFD---FGLRPSV 303
Qy 96 AYDKAASK---THLSASIPDTRMWY-SIGATYKFTPNLSVDVGVGAHLRGKKHFVETQN 151
Db 304 AVLQSKGDISNGYGSYGQDQDIVKYVDVGATYFKNKMSYVDY----- 348
Qy 152 IKGLLIVEADYTTKA---TANLYGLNLNRYF 179
Db 349 -KINLLDKNDFTRDAGINTDDIVALLGLVYQF 378

RESULT 36
OPAG_NEIGO
ID OPAG_NEIGO STANDARD; PRT; 247 AA.
AC Q04875;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Opacity protein OPA52 precursor (OPA31) (Fragment).
GN OPAG.

```


GN VASI OR YGR094W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87222321; PubMed=3294828;
 RA Jordana X., Chation B., Paz-Weisshaar M., Buhler J.-M., Cramer F.,
 RA Ebel J.-P., Fasiolo F.;
 RT "Structure of the yeast valyl-tRNA synthetase gene (VASI) and the
 RT homology of its translated amino acid sequence with Escherichia coli
 RT isoleucyl-tRNA synthetase.";
 RL J. Biol. Chem. 262:7189-7194(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wedler H., Scharif M., Wedler E., Wambutt R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 1-78 FROM N.A.
 RX MEDLINE=88087140; PubMed=3275649;
 RA Chation B., Walter P., Ebel J.-P., Lacroite F., Fasiolo F.;
 RT "The yeast VASI gene encodes both mitochondrial and cytoplasmic
 RT valyl-tRNA synthetases.";
 RL J. Biol. Chem. 263:52-57(1988).
 CC -|- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
 CC + L-valyl-tRNA(Val).
 CC -|- SUBCELLULAR LOCATION: Mitochondrial and cytoplasmic.
 CC -|- ALTERNATIVE PRODUCTS: 2 isoforms; a mitochondrial form (shown
 CC here) and a cytoplasmic form; are produced by alternative
 CC initiation.
 CC -|- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02719; AAA35207.1; -
 DR EMBL; 272879; CAA97057.1; -
 DR EMBL; M18392; AAA35205.1; -
 DR EMBL; M18392; AAA35206.1; -
 DR PIR; A29871; SYBYT.
 DR HSP; P96142; IGAX.
 DR SGD; S0003326; VAS1.
 DR InterPro; IPR002300; tRNA-synt_la.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002303; tRNA-synt_val.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00986; TRNASYNTHAL.
 DR TIGRFAMs; TIGR00422; vals; 1.
 DR PROSITE; PS00178; AA.TRNA.LIGASE.1; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Mitochondrion; Translational peptide; Alternative initiation.
 FT TRANSIT 1 47
 FT CHAIN 48 1104
 FT CHAIN 47 1104
 FT VALYL-TRNA SYNTHETASE, MITOCHONDRIAL
 FT ISOFORM.
 FT VALYL-TRNA SYNTHETASE, CYTOPLASMIC
 FT ISOFORM.
 FT INIT_MET 47
 FT SITE 190 200
 FT "HIGH" REGION.
 FT SITE 703 707
 FT "KMSKS" REGION.
 FT SITE 706 706
 FT ATP (BY SIMILARITY).
 FT CONFLICT 147 147
 FT A -> G (IN REF. 1).
 FT CONFLICT 540 540
 FT R -> K (IN REF. 1).
 FT SEQUENCE 1104 AA; 125769 MW; 6493AEF37ECD4A7C CRC64;
 SQ

Query Match 7.8%; Score 75; DB 1; Length 1104;
 Best Local Similarity 20.7%; Pred. No. 40;
 Matches 36; Conservative 25; Mismatches 63; Indels 50; Gaps 6;
 QY 20 PAYWELSGFHDTDWAIIHYSKYTEWSRFKELRGYQDGSYGAEFTKKEEYKDNRSR--- 76
 DB 812 PATEGLSGNSLEVKWLH---KLTETSKVN-----EALDKRDLFTSTSSYE 857
 QY 77 --PAITTTSLNDALTLRAGLAYDKAASKTHLSASIPD-----TDRMWSIGA 122
 DB 858 FMYLICDVYIENSKYLIQESAIKESAKDTLYLLDNLAKLIHPFMPFISEEMWQRL-- 915
 QY 123 TYKFTPNLSVDVGFALRGKKHVFETQNIKGLLLVADYTTKATKATNLYGLNLN 176
 DB 916 -----PKRSTERKASIVKASYPYV-----SEYDDVKSANAYDLVLN 952
 RESULT 39
 POLG.PEMVM
 ID POLG.PEMVM STANDARD; PRT; 3099 AA.
 AC 056075;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: N-terminal protein (PI); Helper
 DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
 DE protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
 DE (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
 DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
 DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (RC
 DE 2.7.7.48); Coat protein (CP)].
 OS Peanut mottle virus (strain M).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=103926;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Flasinaki S., Gonzales R.A., Cassidy B.G.;
 RT "The complete nucleotide sequence of peanut mottle virus (M strain)
 RT genomic RNA.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -|- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -|- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -|- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
 CC further restricted by preferences for the amino acids in P6 - P1'
 CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 CC Glu+(Ser or Gly) for the enzyme from tobacco etch virus. The
 CC natural substrate is the viral polyprotein, but other proteins and
 CC oligopeptides containing the appropriate consensus sequence are
 CC also cleaved.
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -|- CATALYTIC ACTIVITY: Hydrolyzes a Gly-I-Gly bond at its own C-
 CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-I-Gly, in the
 CC processing of the polyviral polyprotein.
 CC -|- PTH: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -|- PTH: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -|- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -|- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -|- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	580.5	60.7	427	2	Q4R051	haemophilus	
2	571.5	59.7	449	2	Q9K2Q4	haemophilus	
3	571.5	59.7	449	2	Q9KHF9	haemophilus	
4	570.5	59.6	449	2	Q00754	haemophilus	
5	564	58.9	451	2	Q9KHF8	haemophilus	
6	563.5	58.9	451	2	Q9KHF7	haemophilus	
7	562.5	58.8	449	2	Q9KHF5	haemophilus	
8	561	58.6	455	2	Q9KHH2	haemophilus	
9	561	58.6	455	2	Q9KHH1	haemophilus	
10	559.5	58.5	456	2	Q9KHG3	haemophilus	
11	559	58.4	455	2	Q9KHH4	haemophilus	
12	559	58.4	455	2	Q9KHH3	haemophilus	
13	559	58.4	455	2	Q00450	haemophilus	
14	557	58.2	455	2	Q9KHH5	haemophilus	
15	553.5	57.8	451	2	Q9KHf6	haemophilus	
16	552.5	57.7	450	2	Q9KHF4	haemophilus	

Db 32i SNNSRVAIGASYNIDEKI.TI.BAGTAWDCC

Qy	72	KNSRFAICTTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWSIGATYKFTPNLS	131
		: : : : : : : :	
Db	324	SNNRSVALGASYNUDEKLTIRAGIANDQAASHHRSAAIPDTRDTWYSLGATYKFTPNLS	380

AC	O9KJH8:
AD	01-OCT-2000 (TReMBLrel. 15,
AE	DT Created)
AF	01-OCT-2000 (TReMBLrel. 15,
AG	DT Last sequence update)
AH	01-OCT-2000 (TReMBLrel. 15,
AI	DT Last annotation update)
AJ	DE Outer membrane protein pl.
AK	DE OMppl.
AL	GN Haemophilus influenzae
AM	OS

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RP [1]

RC SEQUENCE FROM N.A.

RC STRAIN=2-H-1038;

RX MEDLINE=20359342; PubMed=10899849;

RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;

RT "Variability of outer membrane protein PI and its evaluation as a

RT vaccine candidate against experimental otitis media due to nontypeable

RT Haemophilus influenzae: an unambiguous, multifaceted approach.";

RL Infect. Immun. 68:4505-4517(2000).

DR EMBL; AF260371; AAF97585.1; -

SQ SEQUENCE 451 AA; 48631 MW; E9D7DF993930C0BDB CRC64;

Query Match 58.9%; Score 564; DB 2; Length 451;

Best Local Similarity 60.7%; Pred. No. 3.9e-41;

Matches 108; Conservative 29; Mismatches 37; Indels 4; Gaps 3;

QY 4 GVLGPYIGK-GSLTLKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDCGY 62

DB 276 GKKGIEAGTGNLTFTLPDYLESGFHLTDKFAVHYSKYTHWSRLTKLHASYEDCK-- 333

QY 63 EAFPTKEEYKONSRAFTGTTYSLNDALTLRAGLAYDKAAKTHLSASIPDTRMWSYGA 122

DB 334 KAFDKELQYSNRSRVALGASYNLDKLTLAGIAYDOAASRHRSAAIPDTRTWYSLGA 393

QY 123 TYKFTPNLSVDVGFAHLRGKKHVFETQNI-KGLLVLEADYTTKATANLYGLNLNRYF 179

DB 394 TYKFTPNLSVDLGAYLKGKVKHFKEVQKVAADFTANANTYTSQAHANLYGLNLNYSF 451

RESULT 6

Q9KHF7

ID Q9KHF7 PRELIMINARY; PRT: 451 AA.

AC Q9KHF7;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Outer membrane protein PI.

GN OMPL.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RP [1]

RC SEQUENCE FROM N.A.

RC STRAIN=658;

RX MEDLINE=20359342; PubMed=10899849;

RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;

RT "Variability of outer membrane protein PI and its evaluation as a

RT vaccine candidate against experimental otitis media due to nontypeable

RT Haemophilus influenzae: an unambiguous, multifaceted approach.";

RL Infect. Immun. 68:4505-4517(2000).

DR EMBL; AF260372; AAF97586.1; -

SQ SEQUENCE 451 AA; 48796 MW; 3C6631D057EB0FB CRC64;

Query Match 58.9%; Score 563.5; DB 2; Length 451;

Best Local Similarity 62.9%; Pred. No. 4.3e-41;

Matches 105; Conservative 28; Mismatches 31; Indels 3; Gaps 2;

QY 14 SLTLKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDCGYEATTKKEEYKD 73

DB 287 NLTLTLPDYLESGFHLTDKFAVHYSKYTHWSRLTRYASYEDCK--KVPDKELQYSN 344

QY 74 NSRAFTGTTYSLNDALTLRAGLAYDKAAKTHLSASIPDTRMWSYIGATYKFTPNLSVD 133

DB 345 NSRVALGASYNLDKLTLAGIAYDOAASRHRSAAIPDTRTWYSLGATYKFTPNLSVD 404

QY 134 VGFAHLRGKKHVFETQNI-KGLLVLEADYTTKATANLYGLNLNRYF 179

DB 394 TYKFTPNLSVDLGAYLKGKVKHFKEVQKVAADFTANANTYTSQAHANLYGLNLNYSF 451

DB 405 LGVAYLKGKVKHFKEVQKVAADGLITTNANTYTSQAHANLYGLNLNYSF 451

RESULT 7

Q9KHF5

ID Q9KHF5 PRELIMINARY; PRT: 449 AA.

AC Q9KHF5;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE Outer membrane protein PI.

GN OMPL.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RP [1]

RC SEQUENCE FROM N.A.

RC STRAIN=9;

RX MEDLINE=20359342; PubMed=10899849;

RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;

RT "Variability of outer membrane protein PI and its evaluation as a

RT vaccine candidate against experimental otitis media due to nontypeable

RT Haemophilus influenzae: an unambiguous, multifaceted approach.";

RL Infect. Immun. 68:4505-4517(2000).

DR EMBL; AF260374; AAF97588.1; -

SQ SEQUENCE 449 AA; 48424 MW; F536A974B0C7F827 CRC64;

Query Match 58.8%; Score 562.5; DB 2; Length 449;

Best Local Similarity 60.8%; Pred. No. 5.3e-41;

Matches 107; Conservative 28; Mismatches 38; Indels 3; Gaps 2;

QY 4 GVLGPYIGKSLTLKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDCGYE 63

DB 277 GVIGAG-KKGDLTLPDYLESGFHLTDKLAHVHYSKYTHWSRLTKLANFEDCK--K 333

QY 64 AFTKKEEYKONSRAFTGTTYSLNDALTLRAGLAYDKAAKTHLSASIPDTRMWSYIGAT 123

DB 334 AFDKELQYSNRSRVALGASYNLYEKLTLRAGIAYDOAASCHHRSAAIPDTRTWYSLGAT 393

QY 124 YKFTPNLSVDVGFAHLRGKKHVFETQNIKGLLVLEADYTTKATANLYGLNLNRYF 179

DB 394 YKFTPNLCVDLGAYLKGKVKHFKEVQVGGHIITNANTYTSQAHANLYGLNLNYSF 449

RESULT 8

Q9KHH2

ID Q9KHH2 PRELIMINARY; PRT: 455 AA.

AC Q9KHH2;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE Outer membrane protein PI.

GN OMPL.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RP [1]

RC SEQUENCE FROM N.A.

RC STRAIN=161;

RX MEDLINE=20359342; PubMed=10899849;

RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;

RT "Variability of outer membrane protein PI and its evaluation as a

RT vaccine candidate against experimental otitis media due to nontypeable

RT Haemophilus influenzae: an unambiguous, multifaceted approach.";

RL Infect. Immun. 68:4505-4517(2000).

DR EMBL; AF260353; AAF97567.1; -

SQ SEQUENCE 455 AA; 49139 MW; 9CCB2EB675B421D9 CRC64;

Query Match 58.6%; Score 561; DB 2; Length 455;

```
Best Local Similarity 62.2%; Pred. No. 7.2e-41;
Matches 107; Conservative 30; Mismatches 29; Indels 6; Gaps 3;

QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSYKYTWSEFRKELRGKYQDGSYGAEATKKEEY 71
DB 286 KGNLTLPDYLESLGSGFHLTDKFAVHYSYKYTHWSRLTKLHASFEDGK--KAFDKELQY 343
QY 72 KDNSRFAIGTTSYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWYSIGATYKFTPNLS 131
DB 344 SNNRVALGASYNLDEKLTLAGIAYDOAASRRHRSRAIPTDRTWYSLGATYKFTPNLS 403
QY 132 VDVGFAHLRGKKHVFETQNI---KGLL-VEADYTTKATANLYGLNLYRF 179
DB 404 VDLGAYLKGKKVHFKVEKTKGDKRALTLNTTANYTSOAHANLYGLNLYSF 455

RESULT 9
Q9KHH1 ID Q9KHH1 PRELIMINARY; PRT; 455 AA.
AC Q9KHH1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT-9;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Trung-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260354; AAF97568.1;
SQ SEQUENCE 455 AA; 49123 MW; FD5882979F10EB24 CRC64;

Query Match 58.6%; Score 561; DB 2; Length 455;
Best Local Similarity 62.2%; Pred. No. 7.2e-41;
Matches 107; Conservative 30; Mismatches 29; Indels 6; Gaps 3;

QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSYKYTWSEFRKELRGKYQDGSYGAEATKKEEY 71
DB 286 KGNLTLPDYLESLGSGFHLTDKFAVHYSYKYTHWSRLTKLHASFEDGK--KAFDKELQY 343
QY 72 KDNSRFAIGTTSYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWYSIGATYKFTPNLS 131
DB 344 SNNRVALGASYNLDEKLTLAGIAYDOAASRRHRSRAIPTDRTWYSLGATYKFTPNLS 403
QY 132 VDVGFAHLRGKKHVFETQNI---KGLL-VEADYTTKATANLYGLNLYRF 179
DB 404 VDLGAYLKGKKVHFKVEKTKGDKRALTLNTTANYTSOAHANLYGLNLYSF 455

RESULT 10
Q9KHG3 ID Q9KHG3 PRELIMINARY; PRT; 456 AA.
AC Q9KHG3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=1-H-1085;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Trung-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260364; AAF97578.1;
SQ SEQUENCE 456 AA; 49133 MW; OD6AFC5DFEB35014 CRC64;

Query Match 58.5%; Score 559; DB 2; Length 456;
Best Local Similarity 62.0%; Pred. No. 9.8e-41;
Matches 106; Conservative 28; Mismatches 32; Indels 5; Gaps 2;

QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSYKYTWSEFRKELRGKYQDGSYGAEATKKEEY 71
DB 288 KGDLTLPDYLESLGSGFHLTDKFAVHYSYKYTHWSRLTKLYASYENGK--KAFDKELQY 345
QY 72 KDNSRFAIGTTSYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWYSIGATYKFTPNLS 131
DB 346 SNNRVALGASYNLDEKLTLAGIAYDOAASRRHRSRAIPTDRTWYSLGATYKFTPNLS 405
QY 132 VDVGFAHLRGKKHVFETQNIKGLL---VEADYTTKATANLYGLNLYRF 179
DB 406 VDLGAYLKGKKVHFKVEKTKGDKRALTLNTTANYTSOAHANLYGLNLYSF 456

RESULT 11
Q9KHH4 ID Q9KHH4 PRELIMINARY; PRT; 455 AA.
AC Q9KHH4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6094;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Trung-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260349; AAF97563.1;
SQ SEQUENCE 455 AA; 49111 MW; OCB00B162A6A6CB1 CRC64;

Query Match 58.4%; Score 559; DB 2; Length 455;
Best Local Similarity 61.6%; Pred. No. 1.1e-40;
Matches 106; Conservative 29; Mismatches 31; Indels 6; Gaps 2;

QY 12 KGSITLKLPAWELSGFHLTDQWAIHYSYKYTWSEFRKELRGKYQDGSYGAEATKKEEY 71
DB 286 KGNLTLPDYLESLGSGFHLTDKFAVHYSYKYTHWSRLTKLHASFEDGK--KAFDKELQY 343
QY 72 KDNSRFAIGTTSYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWYSIGATYKFTPNLS 131
DB 344 SNNRVALGASYNLDEKLTLAGIAYDOAASRRHRSRAIPTDRTWYSLGATYKFTPNLS 403
QY 132 VDVGFAHLRGKKHVFETQNIKG---LLLVEADYTTKATANLYGLNLYRF 179
DB 404 VDLGAYLKGKKVHFKVEKTKGDKRALTLNTTANYTSOAHANLYGLNLYSF 455

RESULT 12
```

CC -I- SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC -I- SIMILARITY: STRONG, TO E.COLI LONG-CHAIN FATTY ACID TRANSPORT
CC PROTEIN (FADL).
DR ENBL; M63151: AAA24991.1; -.
DR ENBL; AP260351: AAF97565.1; -.
DR ENBL; AF260350: AAF97564.1; -.
DR PIR; A40183: A40183.
KW Outer membrane; Signal; Transmembrane.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 455 OUTER MEMBRANE PROTEIN P1.
SQ SEQUENCE 455 AA; 49141 MW; 799059130A3D9DE1 CRC64;
Query Match 58.4%; Score 559; DB 2; Length 455;
Best Local Similarity 61.6%; Pred. No. 1.le-40;
Matches 106; Conservative 29; Mismatches 31; Indels 6; Gaps
Qy 12 KGSITLKLPAYWELSGFHQLTDQWAITHYSKYKTEWSRFKELRGKYODGSGYEAFTKKEEY 71
Db 286 KGNLTLPDYLELSGFHQLTDFAVHYSKYKTHWSRLTKLHASPEDGK--KAPDKELQY 343
Qy 72 KDNSRFAIGTYSUNDALTURLAGLYDKAAKSTHLASIPDTDRWMYSIGATYKFTPNLS 131
Db 344 SNNSRVALGASYNDLKLTLRAGIADQAASHRRHSAAIPDTRTWYSLGATYKFTPNLS 403
Qy 132 VDVGFAHLRGKKKHFEVTONIKG----LLLVEADYTTKATANLYGLNLNYRF 179
Db 404 VDLGYALRGKKVHFKEVTIGDKRTLNTNTANTYSQAHANLYGLNLNYSF 455
RESULT 14
QKHH5 PRELIMINARY; PRT; 455 AA.
AC O9KHH5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBurel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBurel. 15, Last annotation update)
DE Outer membrane protein P1.
OS OMPII.
GN Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7004;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR ENBL; AF260348: AAF97562.1; -.
SQ SEQUENCE 455 AA; 49165 MW; 280E21B16CA18C0 CRC64;
Query Match 58.2%; Score 557; DB 2; Length 455;
Best Local Similarity 61.6%; Pred. No. 1.6e-40;
Matches 106; Conservative 28; Mismatches 32; Indels 6; Gaps
Qy 12 KGSITLKLPAYWELSGFHQLTDQWAITHYSKYKTEWSRFKELRGKYODGSGYEAFTKKEEY 71
Db 286 KGNLTLPDYLELSGFHQLTDFAVHYSKYKTHWSRLTKLHASPEDGK--KAPDKELQY 343
Qy 72 KDNSRFAIGTYSUNDALTURLAGLYDKAAKSTHLASIPDTDRWMYSIGATYKFTPNLS 131
Db 344 SNNSRVALGASYNDLKLTLRAGIADQAASHRRHSAAIPDTRTWYSLGATYKFTPNLS 403
Qy 132 VDVGFAHLRGKKKHFEVTONIKG----LLLVEADYTTKATANLYGLNLNYRF 179
Db 404 VDLGYALRGKKVHFKEVTIGDKRTLNTNTANTYSQAHANLYGLNLNYSF 455

Q9KHF6	PRELIMINARY;	PRT;	451 AA.
ID	Q9KHF6		
AC	Q9KHF6;		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TRENBLrel. 15, Last annotation update)		
DE	Outer membrane protein p1.		
GN	OMPp1.		
OS	Haemophilus influenzae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae.		
OC	Haemophilus.		
ON	NCBI_TaxID=727;		
OX	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=13-H-1157;		
RC	MEDLINE=20359342; PubMed=10899849;		
RA	Bolduc G.R., Bouchet V., Jiang R.Z., Geiselsoder J.,		
RA	Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;		
RT	"Variability of outer-membrane protein p1 and its evaluation		
RT	vaccine candidate against experimental otitis media due to no		
RT	Haemophilus influenzae: an unambiguous, multifaceted approach		
RL	Infect. Immun. 68:4505-4517(2000).		
RL	EMBL; AF260373; AAF97587.1; -.		
SQ	SEQUENCE 451 AA; 48446 MW; F4DD7C9354DFCF13 CRC64;		
	Query Match 57.8%; Score 553.5; DB 2; Length 451;		
	Best Local Similarity 62.1%, Pred. No. 3.2e-40;		
	Matches 105; Conservative 27; Mismatches 34; Indels 3		
QY	12 KGSLLTKLPAYWELSGFHQTDQAIHYSYKYTWSRKEKLRGKYQDGSYGAEFTK		
	: : : : : : : : : : : : : : :		
Db	285 KGDLTLTPDYLELSGFHQLDKLAVHYSKYKTHWSRLTKLNASFEDGK --KAFDK		
	: : : : : : : : : : : : : : :		
QY	72 KDNSFAIGTGYSLNDALTRLAGLVAKDAAKTHLSASIPDTRMWSYIGATYKFT		
	: : : : : : : : : : : : : : :		
Db	343 SNNSRVALGASYNLEKLTLAGIAYDQAASRRHSAAIPDTRTWYSLGATYKFT		
	: : : : : : : : : : : : : : :		
QY	132 VDVGFAHLRGKKHVFETQNIK -GLLLVADYTTKATANYGLNLYRF 179		
	: : : : : : : : : : : : : : :		
Db	403 VDLGAYILKGKHFKEVQKAAGGHTITANYTQSAHANYGLNLYSF 451		

```

RESULT 16
Q9KHf4
ID Q9KHf4 PRELIMINARY; PRT; 450 AA.
AC Q9KHf4;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE Outer membrane protein Pl.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7424;
RX MEDLINE=20359342; PubMed=10899849;
RA Boilduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc O.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein Pl and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RT Infect. Immun. 68:4505-4517(2000).
RL EMBL; AF260375; AAF97589.1; -.
DR EMBL; AF260375; AAF97589.1; -.
SQ SEQUENCE 450 AA; 48593 MW; CF17201C42BDA757 CRC64;

Query Match 57.7%; Score 552.5; DB 2; Length 450;
Best Local Similarity 61.8%; Pred. No. 3.9e-40;
Matches 105; Conservative 27; Mismatches 35; Indels 3; Gaps 2;

QY 11 GK-GSLTKLPAYWELSGFHQLTDQIAIHYSKYKTEWSRFKELRGKYQDGSVEAFTKKE 69

```

Db	283	GKTGDTLTTPDYLELSGGFHQTDKLVAVHSYKYTHWSRLTKLHASYENGE--KAFDKEL	340			
Qy	70	EYKDNSRFAIGTYYNSLNDALTLRAGLAYDKAASKTHLSASIPDTRDMWYSIGATKYKFTPN	129			
Db	341	QYSNNSRVALGASYNLYEKLTLRAGLAYDQOASRHQRSAAIPDTRTWYSIGATKYKFTPN	400			
Qy	130	LSVDVGFAHLRGKKKHFEVQTNTKGLLLVEADYTTKATANLYGLNLNRYF	179			
Db	401	LSVDLGAYLKGKKVHKFEKQAGGHIITNANYTSOAHANLYGLNLNYSF	450			
RESULT 17						
Qy	Q9KHF3	PRELIMINARY; PRT; 450 AA.				
AC	O9KHF3					
DT	01-OCT-2000	(TrEMBLrel. 15, Created)				
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)				
DT	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)				
DE	Outer membrane protein P1.					
GN	OMPPI.					
OS	Haemophilus influenzae.					
OC	Bacteria; Proteobacteria; gamma-subdivision; Pasteurellaceae;					
OC	Haemophilus.					
OX	NCBI_TaxID=727;					
RP	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=4-H-1093;					
RX	MEDLINE=20359342; PubMed=10899849;					
RA	Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,					
RA	"Variability of outer membrane protein P1 and its evaluation as a					
RT	vaccine candidate against experimental otitis media due to nontypeable					
RT	Haemophilus influenzae: an unambiguous, multifaceted approach."					
RL	Infect. Immun. 68:4505-4517(2000).					
RL	EMBL; AP260376; AAF97590.1; --					
SQ	SEQUENCE 450 AA; 48741 MW; 3C68E144339F41CD CRC64;					
Query Match						
Best Local Similarity 61.8%; Pred. No. 3.9e-40;						
Matches 105; Conservative 28; Mismatches -34; Indels 3; Gaps						
Qy	11	GK-GSLTKLPATWELSGGFHQLTDQWAIHYSKYTWFSRFEKLRKDYDGGSYEAFTKKE	69			
Db	283	GKTGDTLTTPDYLELSGGFHQTDKLVAVHSYKYTHWSRLTKLHASYENGE--KAFDKEL	340			
Qy	70	EYKDNSRFAIGTYYNSLNDALTLRAGLAYDKAASKTHLSASIPDTRDMWYSIGATKYKFTPN	129			
Db	341	QYSNNSRVALGASYNLYEKLTLRAGLAYDQOASRHQRSAAIPDTRTWYSIGATKYKFTPN	400			
Qy	130	LSVDVGFAHLRGKKKHFEVQTNTKGLLLVEADYTTKATANLYGLNLNRYF	179			
Db	401	LSVDLGAYLKGKKVHKFEKVTIGAHITNANYTSOAHANLYGLNLNYSF	450			

QY 11 GK-GSLTLKLPAYWELSGFHQLTDOAWIHYSYKYTEWSREKELRGKYQDGCSGYEAFTKKE 69

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260367; AAF97581.1; -.
SQ SEQUENCE 451 AA; 48640 MW; 72C6977C6F2A7BA1 CRC64;

Query Match 57.5%; Score 550.5; DB 2; Length 451;
Best Local Similarity 61.2%; Pred. No. 5.9e-40;
Matches 109; Conservative 26; Mismatches 38; Indels 5; Gaps 3;

QY 4 GVLPYIGKSLTKLPAYWELSGFHQLTQDQAIHYSYKYTEWSRPFKELRGKYQDGSYG 63
DB 277 GVIGAG-KKGDLTLPDYLELSGFHQLTKLAVHYSKYTHWSRLTKLNASFEDGK--K 333
QY 64 AFTKKEEYKDNRFAGTGYTSLNDALTLRAGLAYDKAAKSTHLSASIPDTRMWSIGAT 123
DB 334 AFDKELQYSNNSRVALGASVDLYEKLTLRAGIAYDQAASRHSRAIPDTRTWSILGAT 393
QY 124 YKFTPNLSVDVGFALRGKKHFF--VETQNIKGLLVEADYTTKATANLYGLNLYRF 179
DB 394 YKFTPNLSVDLGYALRGKKVHFKEVKTGKTVILNATADYTSQAHANLYGLNLYSF 451

RESULT 19
QYKHG9
ID QYKHG9 PRELIMINARY; PRT; 456 AA.
AC QYKHG9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdlvslon; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RC STRAIN=1071;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260358; AAF97572.1; -.
SQ SEQUENCE 456 AA; 49130 MW; C16D7F3809AC8FA2 CRC64;

Query Match 57.4%; Score 549.5; DB 2; Length 456;
Best Local Similarity 61.3%; Pred. No. 7.3e-40;
Matches 111; Conservative 28; Mismatches 33; Indels 9; Gaps 5;

QY 4 GVLPYIGK--GSLTKLPAYWELSGFHQLTQDQAIHYSYKYTEWSRPFKELRGKYQDGSYG 62
DB 280 GVIG--AGKTNLTLPDYLELSGFHQLTKLAVHYSKYTHWSRLTKLNASFEDGK-- 335
QY 63 EAFTKKEEYKDNRFAGTGYTSLNDALTLRAGLAYDKAAKSTHLSASIPDTRMWSIGA 122
DB 336 KAFDKELQYSNNSRVALGASVLYEKLTLRAGIAYDQAASRHSRAIPDTRTWSILGA 395
QY 123 YKFTPNLSVDVGFALRGKKHFFVETQNI---KGLLVE-ADYTTKATANLYGLNLYR 178
DB 396 YKFTPNLSVDLGYALRGKKVHFKEGKTADKSGPLTSTANTYTSQAHANLYGLNLYS 455

QY 179 F 179
DB 456 F 456

RESULT 20

QYKHG6
ID QYKHG6 PRELIMINARY; PRT; 456 AA.
AC QYKHG6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus Influenzae.
OC Bacteria; Proteobacteria; gamma subdlvslon; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RC STRAIN=7416;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260347; AAF97561.1; -.
SQ SEQUENCE 456 AA; 49370 MW; 0FA50D60942905D0 CRC64;

Query Match 57.4%; Score 549; DB 2; Length 456;
Best Local Similarity 61.6%; Pred. No. 8.1e-40;
Matches 106; Conservative 28; Mismatches 32; Indels 6; Gaps 3;

QY 12 KGSLLTKLPAYWELSGFHQLTQDQAIHYSYKYTEWSRPFKELRGKYQDGSYEAFTKKEEY 71
DB 287 KGSLLTKLPAYWELSGFHQLTQDQAIHYSYKYTEWSRPFKELRGKYQDGSYEAFTKKEEY 344
QY 72 KDNRSFAICTYTSINDALTLRAGLAYDKAAKSTHLSASIPDTRMWSIGATYKFTPNLS 131
DB 345 SNNSRVALGASVLYEKLTLRAGIAYDQAASRHSRAIPDTRTWSILGATYKFTPNLS 404
QY 132 VDVGFAHLRGKKHFFVETQNI---KGLL-VEADYTTKATANLYGLNLYRF 179
DB 405 VDLGYALRGKKVHFKEVKTIGDERSLTNTANTYTSQAHANLYGLNLYSF 456

RESULT 21
QYKHG7
ID QYKHG7 PRELIMINARY; PRT; 456 AA.
AC QYKHG7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus Influenzae.
OC Bacteria; Proteobacteria; gamma subdlvslon; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RC STRAIN=LQ-127;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Gelselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260360; AAF97574.1; -.
SQ SEQUENCE 456 AA; 49220 MW; 7AC1F56499D94C91 CRC64;

Query Match 57.4%; Score 549; DB 2; Length 456;
Best Local Similarity 62.2%; Pred. No. 8.1e-40;
Matches 107; Conservative 25; Mismatches 34; Indels 6; Gaps 3;

QY 12 KGSLLTKLPAYWELSGFHQLTQDQAIHYSYKYTEWSRPFKELRGKYQDGSYEAFTKKEEY 71

```

Db 287 KGDLTLTLPDYLESGFHLTDKLAHVHYSKYTHWSRLTKLHASPEDGK--KAFDKELQY 344
Qy 72 KDNSRFAIGTTYSNDALTLAGLAYDKAASKTHLSASIPDTRDMWYSIGATYKFTPNLS 131
Db 345 SNNSRVALGASYNLYEKLTLRAGIAYDQAASRHSRAAIPDTRTWYSLGATYKFTPNLS 404
Qy 132 VDVGFALHRLGKKKHVFETQNI---KGL-LLVEADYTTKATANLYGLNLYRF 179
Db 405 VDLGVAYLKGKKVHFKEVNTIGKESGLPLTTTANTYTSQAHANLYGLNLYSF 456

RESULT 22
Q9KHG2
ID Q9KHG2 PRELIMINARY; PRT; 451 AA.
AC Q9KHG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein p1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=6181;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260365; AAF97579.1; -
SQ SEQUENCE 451 AA; 48611 MW; B616AAB6FEF0252A CRC64;

```

```

Query Match 57.2%; Score 547; DB 2; Length 451;
Best Local Similarity 62.4%; Pred. No. 1.2e-39;
Matches 106; Conservative 27; Mismatches 33; Indels 4; Gaps 3;

Qy 12 KGSLLTKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDGSYGAEFTKKEEY 71
Db 284 KGDLTLTLPDYLESGFHLTDKLAHVHYSKYTHWSRLTKLHASPEDGK--KAFDKELQY 341
Qy 72 KDNSRFAIGTTYSNDALTLAGLAYDKAASKTHLSASIPDTRDMWYSIGATYKFTPNLS 131
Db 342 SNNSRVALGASYNLYEKLTLRAGIAYDQAASRHSRAAIPDTRTWYSLGATYKFTPNLS 401
Qy 132 VDVGFALHRLGKKKHVFETQNI---KGLLL-VEADYTTKATANLYGLNLYRF 179
Db 402 VDLGVAYLKGKKVHFKEVATIGKSVILNTTANTYTSQAHANLYGLNLYSF 451

```

```

RESULT 23
Q9K324
ID Q9K324 PRELIMINARY; PRT; 459 AA.
AC Q9K324;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein p1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=218, AND 200;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,

```

```

RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260344; AAF97558.1; -
DR EMBL; AF260343; AAF97557.1; -
SQ SEQUENCE 459 AA; 49717 MW; 54490C08E4729870 CRC64;

Query Match 57.2%; Score 547; DB 2; Length 459;
Best Local Similarity 61.6%; Pred. No. 1.2e-39;
Matches 106; Conservative 27; Mismatches 33; Indels 6; Gaps 3;

Qy 12 KGSLLTKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDGSYGAEFTKKEEY 71
Db 290 KGDLTLTLPDYLESGFHLTDKLAHVHYSKYTHWSRLTKLHASPEDGK--KAFDKELQY 347
Qy 72 KDNSRFAIGTTYSNDALTLAGLAYDKAASKTHLSASIPDTRDMWYSIGATYKFTPNLS 131
Db 348 SNNSRVALGASYNLYEKLTLRAGIAYDQAASRHSRAAIPDTRTWYSLGATYKFTPNLS 407
Qy 132 VDVGFALHRLGKKKHVFETQNI---KGLLL-VEADYTTKATANLYGLNLYRF 179
Db 408 VDLGVAYLKGKKVHFKEVKTIGDERSALNTTANTYTSQAHANLYGLNLYSF 459

```

```

RESULT 24
Q9KHH8
ID Q9KHH8 PRELIMINARY; PRT; 459 AA.
AC Q9KHH8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein p1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=667;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260345; AAF97559.1; -
SQ SEQUENCE 459 AA; 49773 MW; 5E6DCD78F9A8FF8A CRC64;

```

```

Query Match 57.2%; Score 547; DB 2; Length 459;
Best Local Similarity 61.6%; Pred. No. 1.2e-39;
Matches 106; Conservative 27; Mismatches 33; Indels 6; Gaps 3;

Qy 12 KGSLLTKLPAYWELSGFHLTDQWAIHYSKYTWSRFKELRGKYQDGSYGAEFTKKEEY 71
Db 290 KGDLTLTLPDYLESGFHLTDKLAHVHYSKYTHWSRLTKLHASPEDGK--KAFDKELQY 347
Qy 72 KDNSRFAIGTTYSNDALTLAGLAYDKAASKTHLSASIPDTRDMWYSIGATYKFTPNLS 131
Db 348 SNNSRVALGASYNLYEKLTLRAGIAYDQAASRHSRAAIPDTRTWYSLGATYKFTPNLS 407
Qy 132 VDVGFALHRLGKKKHVFETQNI---KGLLL-VEADYTTKATANLYGLNLYRF 179
Db 408 VDLGVAYLKGKKVHFKEVKTIGDERSALNTTANTYTSQAHANLYGLNLYSF 459

RESULT 25
Q9KHH7
ID Q9KHH7 PRELIMINARY; PRT; 459 AA.
AC Q9KHH7;

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein p1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=199;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260346; AAF97560.1; -.
SQ SEQUENCE 459 AA; 49745 MW; 5A46FF07187E98B3 CRC64;

Query Match 57.2%; Score 547; DB 2; Length 459;
Best Local Similarity 61.6%; Pred. No. 1.2c-39;
Matches 106; Conservative 27; Mismatches 33; Indels 6; Gaps 3;

QY 12 KGSLLKLPAYWELSGFHQLTDMWAIHYSKYTWSRFKELRGKYQDGSYEAFTKKEEY 71
DB 290 KGDLTLPDYLESGFHQLTDMWAIHYSKYTWSRFLTKLNASFEDGK--KAFDKELQY 347

QY 72 KNSRFAIGTTYSNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYGATYKFTPNLS 131
DB 348 SNNSRVALGASYNLYEKLTRAGIAYDQAASRHSAAIPDTRTWYSLGATYKFTPNLS 407

QY 132 VDVGFAHLRGKKHVFETONI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
DB 408 VDLGYALYLGKKVHFKEVKTIGDRSLALNTTANTYTSQAHANLYGLNLNYSF 459

RESULT 26
Q9KHG8
ID Q9KHG8 PRELIMINARY; PRT; 456 AA.
AC Q9KHG8:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein p1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCH-2;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260359; AAF97573.1; -.
SQ SEQUENCE 456 AA; 49075 MW; CF7B3ECE4213C841 CRC64;

Query Match 57.0%; Score 545.5; DB 2; Length 456;
Best Local Similarity 60.2%; Pred. No. 1.6e-39;
Matches 109; Conservative 29; Mismatches 34; Indels 9; Gaps 5;

QY 4 GVLGPVIGK-GSLTKLPAYWELSGFHQLTDMWAIHYSKYTWSRFKELRGKYQDGSY 62
DB 280 GVIG--AGTGNLTFTLPDYLESGFHQLTDMWAIHYSKYTWSRFLTKLNASFEDGK-- 335

```

```

QY 63 EAFTKKEEYKDNSRFAIGTTYSNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYGA 122
DB 336 KAFDKELQYSNNSRVALGASYNLYEKLTRAGIAYDQAASRHSAAIPDTRTWYSLGA 395

QY 123 TYKFTPNLSVDVGFAHLRGKKHVFETONI---KGL-LLVEADYTTKATANLYGLNLNRY 178
DB 396 TYKFTPNLSVDVGAYLKGKKVHFKEVKTIGKSGLPVTATANTYTSQAHANLYGLNLNYS 455

QY 179 F 179
DB 456 F 456

RESULT 27
Q9KH14
ID Q9KH14 PRELIMINARY; PRT; 456 AA.
AC Q9KH14:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein p1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCH-1;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260337; AAF97551.1; -.
SQ SEQUENCE 456 AA; 49402 MW; 8CF3E87CB81DD6E6 CRC64;

Query Match 56.8%; Score 544; DB 2; Length 456;
Best Local Similarity 61.4%; Pred. No. 2.2e-39;
Matches 105; Conservative 28; Mismatches 32; Indels 6; Gaps 3;

QY 13 GSLLKLPAYWELSGFHQLTDMWAIHYSKYTWSRFKELRGKYQDGSYEAFTKKEEYK 72
DB 288 GDLTLPDYLESGFHQLTDMWAIHYSKYTWSRFLTKLNASFEDGK--KAFDKELQYS 345

QY 73 DNSRFAIGTTYSNDALTLRAGLAYDKAASKTHLSASIPDTRMWSYGATYKFTPNLSV 132
DB 346 NNSRVALGASYNLYEKLTRAGIAYDQAASRHSAAIPDTRTWYSLGATYKFTPNLSV 405

QY 133 DVGFAHLRGKKHVFETONI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
DB 406 DLGYALYLGKKVHFKEVKTIGDRSLALNTTANTYTSQAHANLYGLNLNYSF 456

RESULT 28
Q9KH12
ID Q9KH12 PRELIMINARY; PRT; 456 AA.
AC Q9KH12:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein p1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3-H-1552;
RX MEDLINE=20359342; PubMed=10899849;

```

RA	Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.J., Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.; RT "Variability of outer membrane protein p1 and its evaluation as a RT vaccine candidate against experimental otitis media due to nontypeable RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL	Infect. Immun. 68:4505-4517(2000).
RQ	EMBL; AF260339; KAF97553.1; -- F95A977CBCEAF057 CRC64;
SQ	SEQUENCE 456 AA: 49403 MW: F95A977CBCEAF057 CRC64;
	Query Match 56.8%; Score 544; DB 2; Length 456; Best Local Similarity 61.4%; Pred. No. 2.2e-39; Matches 105; Conservative 28; Mismatches 32; Indels 6; Gaps
QY	13 GSITLKLPAWELSGFHQLTDQWATHYSKYKTEWSRFKELRGKYQDGSYGVEATFKKEEYK 72 : : :
Db	288 GDUTLTPDYLELSGFHQLTDLVAVHYSKYTHWSRSLTKLNASEFDGK -- KAFDKELQYS 345 : : :
QY	73 DNSRFAIGTYTSINDALTLRAGLAYDKAASTHLSSASIPDDTRDWYSIGATYFTPNLSV 132 : : :
Db	346 NNSRVALGASYNLDEKLTLRAGIAYDQAASRHQSAAIPDDRTWYSIGATYFTPNLSV 405 : : :
QY	133 DVGFAHLRGKKKHVFETQNI --- KGLLI - VEADYTTKATANLYGLNLYNRF 179 : :
Db	406 DLGYAYLGKGKVHFKEVTIGDKRSLLASTANTYSQAHANLYGLNLYNRF 456 : :

```

RESULT 29
Q9KHG6
ID Q9KHG6 PRELIMINARY; PRT; 457 AA.
AC Q9KHG6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE outer membrane protein pI.
GN ompI.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=5;
RX MEDLINE=20359342; Pubmed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RA "Variability of outer membrane protein pI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae; an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
RL ENBL; AF260361; AAF97575.1; -.
SQ SEQUENCE 457 AA; 49292 MW; D25C84D6E81A965F CRC64;

```

RESULT 30
Q9KHH9
ID Q9KHH9
AC Q9KHH9;
PRELIMINARY;
PRT; 456 AA.

[illegible]

RESULT 34

Q9KHG4
ID Q9KHG4 PRELIMINARY; PRT; 456 AA.

RA MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Trung-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.,
RT "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT *Haemophilus influenzae*: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
RL EMBL: AF260363; AAF97577.1; -.
SQ SEQUENCE 456 AA, 49253 MW; A3B8BB8891302566 CRC64;
SQ

Qy	12	KGSLTLKPAYWBLSGFHQLTDQWAHHYSYKYTETWSEFRKELRGKYQDGSGYEAFTRKEEY	71
Db	288	KGDLTLTLPDYLELBSGFHQLTDKLAVHYSYKYTHWSRLTFLYASSENGK--KAFDKELQY	345
Qy	72	KDNSRFAIGTYSLNDALTLRAGLAVDYDKAASKTHLSASIPDTRDMYVSGATYKFTPNLS	131
Db	346	SNNSRVALGASYNLDEKLTLRAGIAYDQAASRHQRSAAPDTRNRTWYSLGATYKFTPNLS	405
Qy	132	VDVGFAHLRCKKKHVEFTQNI-KGLLIVE--ADYTTKATANLXGLNLNRYF	179
Db	406	VDLGYAYLGGKKVHFKEVKAITKSVLTDTTANTTQAHANLXGLNLNYSF	456
RESULT 35			
Q9KHG1			
ID	Q9KHG1	PRELIMINARY;	PRT; 454 AA.
AC	Q9KHG1;		
DT	01-OCT-2000	(TEMBLrel. 15, Created)	
DT	01-OCT-2000	(TEMBLrel. 15, Last sequence update)	
DT	01-DEC-2001	(TEMBLrel. 19, Last annotation update)	

```
DE Outer membrane protein p1.
GN OMPp1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22-H-1154; PubMed=10899849;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RA "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260366; AAF97580.1; -
SQ SEQUENCE 454 AA; 49986 MW; 494FBAC2600DA2C8 CRC64;

Query Match 55.7%; Score 533; DB 2; Length 454;
Best Local Similarity 61.0%; Pred. No. 2e-38;
Matches 105; Conservative 26; Mismatches 35; Indels 6; Gaps 3;

Qy 12 KGSITLKLPAWELSGFHQLTDQWAIHYSYKYTWRSFKELRGKYQDGSYGAEATKKEEY 71
Db 285 KGDTLTLPDYLESGFHQLTKLAVHYSYKYTHWSRLTRLYASSENGK--KAFDKEQLQY 342
Qy 72 KDNSRFAIGTYSLNDALTLAGLAYDKAASKTHLSASIPDTRMWSIGATYKFTPNLS 131
Db 343 SNNRVALGASYNLYEKLTLRAGIAYDQASRHORSRAAIPDTRTWYSLGATYKFTPNLS 402
Qy 132 VDVGFAHLRGKKKHFEVETQNI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
Db 403 VDLGAYLKGKKVHFKEVKTIGDARSALNTTANYTSQAHANLYGLNLNYSF 454

RESULT 36
QKHHO ID QKHHO PRELIMINARY; PRT; 456 AA.
AC QKHHO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Outer membrane protein p1.
GN OMPp1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17-H-783; PubMed=10899849;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RA "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260340; AAF97554.1; -
SQ SEQUENCE 456 AA; 49308 MW; CECC10E0B6781A8 CRC64;

Query Match 55.7%; Score 533; DB 2; Length 456;
Best Local Similarity 61.0%; Pred. No. 2e-38;
Matches 105; Conservative 26; Mismatches 35; Indels 6; Gaps 3;

Qy 12 KGSITLKLPAWELSGFHQLTDQWAIHYSYKYTWRSFKELRGKYQDGSYGAEATKKEEY 71
Db 287 KGBLTLPDYLESGFHQLTKLAVHYSYKYTHWSRLTRLYASSENGK--KAFDKEQLQY 344
Qy 72 KDNSRFAIGTYSLNDALTLAGLAYDKAASKTHLSASIPDTRMWSIGATYKFTPNLS 131
Db 345 SNNRVALGASYNLYEKLTLRAGIAYDQASRHORSRAAIPDTRTWYSLGATYKFTPNLS 404
```

```
Qy 132 VDVGFAHLRGKKKHFEVETQNI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
Db 405 VDLGAYLKGKKVHFKEVKTIGDARSALNTTANYTSQAHANLYGLNLNYSF 456

RESULT 37
QKHHO ID QKHHO PRELIMINARY; PRT; 460 AA.
AC QKHHO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein p1.
GN OMPp1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9-H-1194; PubMed=10899849;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RA "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260357; AAF97571.1; -
SQ SEQUENCE 460 AA; 49719 MW; FICEBEB597A5CBA7 CRC64;

Query Match 55.6%; Score 532; DB 2; Length 460;
Best Local Similarity 59.9%; Pred. No. 2.5e-38;
Matches 103; Conservative 27; Mismatches 36; Indels 6; Gaps 3;

Qy 12 KGSITLKLPAWELSGFHQLTDQWAIHYSYKYTWRSFKELRGKYQDGSYGAEATKKEEY 71
Db 291 KADLTFTLPDYLESGFHQLTKLAVHYSYKYTHWSRLTKLHASFEDGK--KAFDKEQLQY 348
Qy 72 KDNSRFAIGTYSLNDALTLAGLAYDKAASKTHLSASIPDTRMWSIGATYKFTPNLS 131
Db 349 SNNRVALGASYNLYEKLTLRAGIAYDQASRHORSRAAIPDTRTWYSLGATYKFTPNLS 408
Qy 132 VDVGFAHLRGKKKHFEVETQNI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
Db 409 VDLGAYLKGKKVHFKEVKTIGEQRSLTFTDTTANYTSQAHANLYGLNLNYSF 460

RESULT 38
QKHHO ID QKHHO PRELIMINARY; PRT; 456 AA.
AC QKHHO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein p1.
GN OMPp1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21-H-1328; PubMed=10899849;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RA "Variability of outer membrane protein p1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260341; AAF97555.1; -
```

```
SQ SEQUENCE 456 AA; 49237 MW; 94A04FOA9A649750 CRC64;
Query Match 54.0%; Score 517; DB 2; Length 456;
Best Local Similarity 60.5%; Pred. No. 5e-37;
Matches 104; Conservative 24; Mismatches 38; Indels 6; Gaps 3;

QY 12 KGSLLTKLPAYWELSGFHQTLQDWAHYSKYTWSRFEKLRGKYQDGSYEAFTKKEEY 71
   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 KGDLTLPDYLELSGFHQLTKLAVHYSKYTHWSRLTRLYASSENGK--KAFDRKELQY 344

QY 72 KDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRDMWYSIGATYKFTPNLS 131
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 345 SNNRSVALGASNYLYEKLTLRAGTAKDQAASCLQRSAAIPDTRTWYSLGATYKFTPNLS 404

QY 132 VDVGFAHLRCKKKHFEVTQNI---KGLLL-VEADYTTKATANLYGLNLNRYF 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 405 VDLGYAYLKKKKVHKFEVKVTIGDARSALNTTANYTFOAHANLYGLNLNRYF 456

RESULT 39
Q9KL84 PRELIMINARY; PRT; 412 AA.
AC Q9KL84;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DE 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Long-chain fatty acid transport protein.
GN VCA0862.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
  Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
  Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
  Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
  McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
  Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
  Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
  cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004414; AAF96760.1;
DR TIGR; VCA0862;
DR InterPro; IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 412 AA; 44784 MW; FB606C929838D891 CRC64;

Query Match 44.0%; Score 421; DB 16; Length 412;
Best Local Similarity 46.9%; Pred. No. 1.1e-28;
Matches 83; Conservative 27; Mismatches 61; Indels 6; Gaps 2;

QY 3 NGVLGPYIGKSLTKLPAYWELSGFHQTLQDWAHYSKYTWSRFEKLRGKYQDGSY 62
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 241 SGIATSAVVPQQLKLPALWELSGFHQLNEQWAVHYSYQTDKSSFEELTATSSQCKNG 300

QY 63 EATKKEEYKDNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRDMWYSIGA 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TCFQKIEQYQDNGRWSYGVATYTLNTTTLRAGLAFDRQAGKATL--SIPDSDRFWYSAGL 358

QY 123 TYKFTPNLSVDVGFAHLRCKKKHFEVTQNIKGLLLVEADYTTKATANLYGLNLNRYF 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 TYTLTPQMTDAGFALVQSRKSGSTEKNQIQGNLQFESE----AVAYISALQFNRYF 411

RESULT 40
Q82D48 PRELIMINARY; PRT; 423 AA.
ID Q82D48
AC Q82D48;
```

```
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative long-chain fatty acid transport protein.
GN FADL OR YPO2744.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Yersinia.
OX NCBI_TaxID=632;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
  Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
  Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
  Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
  Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlysheva I.V.,
  Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
  Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414153; CAC92983.1;
DR InterPro; IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 423 AA; 45668 MW; 8D11F6495FDB9FE9 CRC64;

Query Match 38.9%; Score 372.5; DB 16; Length 423;
Best Local Similarity 44.6%; Pred. No. 1.9e-24;
Matches 79; Conservative 23; Mismatches 70; Indels 5; Gaps 3;

QY 4 GVLGPYIGKSLTKLPAYWELSGFHQTLQDWAHYSKYTWSRFEKLRGKYQDGSY 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 GGLGKTVPGSLTLNLPVWELSGYKNAQWALHYSMAVTWSSFKELKATASNGD--V 308

QY 64 AFTKKEEYKDNSRFAIGTYSLNDALTLRAGLAYDKAASKT-HLSASIPDTRDMWYSIGA 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 LFDKHEGFRDAYRIALGTYTYDDNWTFTGTGIAFDSPIPAGNRSISIPDQDFWLSAGT 368

QY 123 TYKFTPNLSVDVGFAHLRCKKKHFEVTQNIKGLLLVEADYTTKATANLYGLNLNRYF 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 TYAFNKNASVDVGIAVMKQGNVSITKTPAPSNPTYE--FNSKGSAMLYGVNENYTF 423

Search completed: May 12, 2003, 09:41:14
Job time: 32 secs
```


GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:41:45 ; Search time 36 Seconds
(without alignments)
662.552 Million cell updates/sec

Title: US-09-995-493-52
Perfect score: 179
Sequence: 1 QHNGVLGPYIGKSLTKLP.....ADYTTKATANLYGLNLNRYR 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minlimum DB seq length: 0

Maxlimum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Prod. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysls of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	7.3	77	AA12456	H.influenzae stral
2	10	5.6	35	AA40060	Hib OMP p1 peptide
3	10	5.6	36	AA12455	HIBP1-14(400-437)
4	10	5.6	36	AA40066	Hib OMP p1 peptide
5	10	5.6	39	AA12454	HIBP1-13(400-437)
6	10	5.6	39	AA40065	Hib OMP p1 peptide
7	10	5.6	53	AA40100	Hib OMP p1-P2 hybr
8	10	5.6	54	AA40101	Hib OMP p1-P2 hybr
9	10	5.6	55	AA40099	Hib OMP p1-P2 hybr
10	8	4.5	636	AA89665	Putative P. abyssi

11	7	3.9	61	20	AA933535	Tobacco DNA bindin
12	7	3.9	68	20	AA931478	A. thaliana RAP2.4
13	7	3.9	68	21	AA825837	AP2 domain amino a
14	7	3.9	69	20	AA931468	N. tabacum EREBP-2
15	7	3.9	69	21	AA825827	AP2 domain amino a
16	7	3.9	79	21	AA932799	zea mays protein f
17	7	3.9	93	22	AA92876	C glutamicum prote
18	7	3.9	95	21	AA932798	zea mays protein f
19	7	3.9	109	22	AAU54058	Propionibacterium
20	7	3.9	113	22	AAU48379	Propionibacterium
21	7	3.9	127	23	ABP25758	Streptococcus poly
22	7	3.9	134	21	AA947231	Arabidopsis thalia
23	7	3.9	206	19	AAW60957	Streptococcus pneu
24	7	3.9	207	23	ABP26302	Streptococcus poly
25	7	3.9	218	22	AAE02538	A. thaliana transc
26	7	3.9	219	21	AA832762	Eucalyptus grandis
27	7	3.9	230	20	AA921977	Sark protein fragm
28	7	3.9	313	21	AA938454	Arabidopsis thalia
29	7	3.9	313	21	AA938787	Arabidopsis thalia
30	7	3.9	330	21	AA938453	Arabidopsis thalia
31	7	3.9	330	21	AA938786	Arabidopsis thalia
32	7	3.9	334	21	AA938452	Arabidopsis thalia
33	7	3.9	334	21	AA938785	Arabidopsis thalia
34	7	3.9	338	22	AA971496	Human oilfactory re
35	7	3.9	338	22	AA972390	Human OR-like poly
36	7	3.9	466	18	AAW17820	Pseudomonas putida
37	7	3.9	511	18	AAW55454	H. pylori ORF 02ge
38	7	3.9	511	19	AAW98255	H. pylori GHPO 155
39	7	3.9	511	19	AAW71519	Helicobacter polyp
40	7	3.9	511	20	AAW17176	H. pylori outer me
41	7	3.9	511	22	AA846315	H. pylori HPS143 p
42	7	3.9	511	22	AA846364	H. pylori HPS143 p
43	7	3.9	523	18	AAW5286	H. pylori ORF 13ee
44	7	3.9	795	22	AA90056	C glutamicum prote
45	7	3.9	865	22	AA988884	E. coli growth and
46	7	3.9	904	20	AAW20767	Senescence-associa
47	6	3.4	8	19	AAW80767	Bloactive fragment
48	6	3.4	15	14	AA42704	Human TGase3 50 kD
49	6	3.4	15	14	AA42705	Murine TGase3 50 k
50	6	3.4	15	16	AA85292	Japanese cedar pol
51	6	3.4	15	19	AAW57752	Residues 146-160 o
52	6	3.4	15	19	AAW57753	Residues 151-165 o
53	6	3.4	18	20	AA931457	A. thaliana AP2 do
54	6	3.4	18	21	AA825816	AP2 repeat AP2-R2
55	6	3.4	20	15	AA845556	Cry j I pollen all
56	6	3.4	20	15	AA845557	Cry j I pollen all
57	6	3.4	20	16	AA882505	Cry j I Japanese C
58	6	3.4	20	16	AA882506	Cry j I Japanese C
59	6	3.4	20	19	AAW82575	Seq ID 10 from WO
60	6	3.4	22	18	AAW08416	Exosite III from h
61	6	3.4	24	20	AA936717	Fragment of human
62	6	3.4	26	18	AA928439	Erythropoietin rec
63	6	3.4	30	15	AA845583	Cry j I pollen all
64	6	3.4	31	14	AA839868	C peptide RV-C1, r
65	6	3.4	33	12	AA812452	HIBP1-10(339-370)
66	6	3.4	33	14	AA840062	Hib OMP p1 peptide
67	6	3.4	34	22	AB830954	Peptide #3605 enco
68	6	3.4	34	22	AB836145	Peptide #3651 enco
69	6	3.4	34	22	AB821531	Protein #3530 enco
70	6	3.4	34	22	AAW69299	Human bone marrow
71	6	3.4	34	22	AAW69315	Peptide #3577 enco
72	6	3.4	34	22	AAW17143	Peptide #3673 enco
73	6	3.4	34	22	AAW29636	Peptide #3520 enco
74	6	3.4	34	22	AAW04838	Human peptide enco
75	6	3.4	34	23	ABG38931	Peptide #7970 enco
76	6	3.4	34	22	AB840464	Protein #6808 enco
77	6	3.4	37	22	AB824809	Human brain expres
78	6	3.4	37	22	AA661286	Human bone marrow
79	6	3.4	37	22	AAW74012	Peptide #6628 enco
80	6	3.4	37	22	AAW20194	Peptide #8221 enco
81	6	3.4	37	22	AAW34184	Human peptide enco
82	6	3.4	37	23	ABG43907	Human neurofilamen
83	6	3.4	44	19	AA920741	

84	6	3.4	49	21	AA056775	Arabidopsis thaliana	157	6	3.4	96	22	AAU67835	Propionibacterium
85	6	3.4	50	23	ABP06117	Human ORFX protein	158	6	3.4	98	22	AAU62250	Propionibacterium
86	6	3.4	53	20	ABY31462	A. thaliana RAP2.7	159	6	3.4	98	22	AAW91797	Human immune/haema
87	6	3.4	53	21	AAE25821	RAP2.7-R2 direct r	160	6	3.4	99	21	AAE40876	Human ORFX ORF640
88	6	3.4	54	23	AAE23550	Human FAIL Ig doma	161	6	3.4	99	23	ABP06513	Human ORFX protein
89	6	3.4	55	20	AAE12333	Human 5' ESR secr	162	6	3.4	100	22	ABG00419	Novel human diagno
90	6	3.4	55	21	AAE56774	Arabidopsis thaliana	163	6	3.4	100	22	AAU31707	Novel human secret
91	6	3.4	55	21	AAE03028	Human secreted pro	164	6	3.4	101	22	AAU00797	Human secreted pro
92	6	3.4	58	22	AAU56725	Propionibacterium	165	6	3.4	104	22	ABG29606	Novel human diagno
93	6	3.4	53	21	AAE56335	Arabidopsis thaliana	166	6	3.4	104	22	ABE43053	Peptide #10559 enc
94	6	3.4	63	21	AAE58828	Arabidopsis thaliana	167	6	3.4	104	22	ABE26224	Protein #8223 enc
95	6	3.4	64	22	AAU58118	Propionibacterium	168	6	3.4	104	22	AAW63955	Human brain expres
96	6	3.4	64	22	ABE28986	Peptide #1637 enc	169	6	3.4	104	22	AAW76775	Human bone marrow
97	6	3.4	64	22	ABE34154	Peptide #1637 enc	170	6	3.4	104	22	AAW20996	Peptide #7430 enc
98	6	3.4	64	22	ABE19598	Protein #1597 enc	171	6	3.4	104	22	AAW36882	Peptide #10919 enc
99	6	3.4	64	22	AAW54946	Human brain expres	172	6	3.4	104	22	ABG45940	Human peptide enc
100	6	3.4	64	22	AAW67326	Human bone marrow	173	6	3.4	104	22	ABG02829	Novel human diagno
101	6	3.4	64	22	AAW15161	Peptide #1595 enc	174	6	3.4	105	22	ABO13822	Human polypeptide
102	6	3.4	64	22	AAW27618	Peptide #1655 enc	175	6	3.4	106	22	AAO02661	Human polypeptide
103	6	3.4	64	22	AAW02899	Peptide #1581 enc	176	6	3.4	107	23	ABE97814	Human secretory po
104	6	3.4	64	23	ABG36969	Human peptide enc	177	6	3.4	109	21	AAV75634	Neisseria gonorrhoe
105	6	3.4	66	22	AAW89590	Human immune/haema	178	6	3.4	109	22	ABG28332	Novel human diagno
106	6	3.4	68	20	AAV31454	A. thaliana AP2 do	179	6	3.4	110	22	AAU86779	Novel human connec
107	6	3.4	68	21	ABE25814	AP2 direct repeat	180	6	3.4	112	11	AAU05219	GX3262 antigen enc
108	6	3.4	68	22	ABE43752	Peptide #11238 enc	181	6	3.4	112	22	AAU50091	Propionibacterium
109	6	3.4	68	22	ABE43924	Peptide #11430 enc	182	6	3.4	113	21	AAO03930	Human secreted pro
110	6	3.4	68	22	ABE16899	Human nervous syst	183	6	3.4	114	20	AAW97370	Human 1-40 OH grow
111	6	3.4	68	22	ABE26678	Protein #8677 enc	184	6	3.4	116	18	AAW28116	Amino acid sequenc
112	6	3.4	68	22	AAW64694	Human brain expres	185	6	3.4	119	19	AAW98819	Eucalyptus grandis
113	6	3.4	68	22	AAW64914	Human brain expres	186	6	3.4	122	21	ABE33140	Human nervous syst
114	6	3.4	68	22	AAW77469	Human bone marrow	187	6	3.4	122	22	ABE17626	Human polypeptide
115	6	3.4	68	22	AAW77650	Human bone marrow	188	6	3.4	122	22	AAO04820	Human polypeptide
116	6	3.4	68	22	AAW21395	Peptide #7829 enc	189	6	3.4	122	22	AAE76439	Human colon cancer
117	6	3.4	68	22	AAW21571	Peptide #8005 enc	190	6	3.4	123	21	AAO3927	Human secreted pro
118	6	3.4	68	22	AAW37630	Peptide #11667 enc	191	6	3.4	124	22	AAO05396	Human polypeptide
119	6	3.4	68	22	AAW37850	Peptide #11887 enc	192	6	3.4	125	20	AAW97365	Bovine growth horm
120	6	3.4	68	23	ABG46487	Human peptide enc	193	6	3.4	125	20	AAW97366	Porcine growth hor
121	6	3.4	68	23	ABG46682	Human peptide enc	194	6	3.4	125	20	AAW97367	Ovine growth hormo
122	6	3.4	69	20	AAV31467	N. tabacum EREBP-1	195	6	3.4	125	20	AAW97369	Caprine growth hor
123	6	3.4	69	21	AAE56970	Human prostate can	196	6	3.4	126	18	AAW27648	Secreted protein B
124	6	3.4	69	21	AAE25826	AP2 domain amino a	197	6	3.4	126	18	AAW44084	Human secreted pro
125	6	3.4	69	22	AAW89883	Human immune/haema	198	6	3.4	126	21	AAE10240	Human adult ovary
126	6	3.4	69	22	AAE73849	Human colon cancer	199	6	3.4	131	20	AAV73843	Human prostate tum
127	6	3.4	69	23	ABP10635	Human ORFX protein	200	6	3.4	131	22	ABG13883	Novel human diagno
128	6	3.4	71	21	AAO03201	Human secreted pro	201	6	3.4	132	22	ABG27033	Novel human diagno
129	6	3.4	72	22	AAU39221	Propionibacterium	202	6	3.4	132	22	ABG90564	C glutamic prote
130	6	3.4	72	23	ABP08259	Human ORFX protein	203	6	3.4	132	23	ABG60172	Human DTHP polye
131	6	3.4	76	22	AAU18559	Human lung antigen	204	6	3.4	133	22	AAE62434	S. epidermidis ope
132	6	3.4	77	23	ABP33821	Human hydrolase-11	205	6	3.4	135	22	AAO04104	Human polypeptide
133	6	3.4	77	23	ABE55088	Lactococcus lactis	206	6	3.4	136	18	AAW24536	Immunoglobulin rB6
134	6	3.4	78	23	ABP26123	Streptococcus poly	207	6	3.4	136	22	ABG28271	Novel human diagno
135	6	3.4	78	23	ABP26124	Streptococcus poly	208	6	3.4	137	22	ABO09065	Thermus caldophilu
136	6	3.4	79	23	ABP02069	Human ORFX protein	209	6	3.4	137	23	AAU80552	Human G protein-co
137	6	3.4	80	23	AAU83080	Novel secreted pro	210	6	3.4	141	21	AAE40829	Human ORFX ORF593
138	6	3.4	82	20	AAV76592	Human ovarian tumo	211	6	3.4	141	23	ABP03564	Breast and ovarian
139	6	3.4	82	22	AAE24466	C glutamic prote	212	6	3.4	142	21	AAE59091	Human secreted pro
140	6	3.4	82	23	ABP27759	Streptococcus poly	213	6	3.4	142	22	ABG27260	Arabidopsis thaliana
141	6	3.4	85	22	AAU46894	Propionibacterium	214	6	3.4	142	23	ABG60149	Human DTHP polye
142	6	3.4	85	22	AAO11082	Human polypeptide	215	6	3.4	145	21	AAE61604	Arabidopsis thaliana
143	6	3.4	88	22	ABE17025	Human nervous syst	216	6	3.4	147	18	AAW28324	Staphylococcus aur
144	6	3.4	90	21	AAV75956	Rat skin cell prot	217	6	3.4	149	20	AAO26996	Human secreted pro
145	6	3.4	90	22	ABG26441	Novel human diagno	218	6	3.4	152	21	AAO59317	Arabidopsis thaliana
146	6	3.4	90	22	ABE38706	Peptide #6212 enc	219	6	3.4	153	21	AAE57110	Human prostate can
147	6	3.4	90	22	AAW59339	Human brain expres	220	6	3.4	155	21	AAE33141	Pinus radiata tran
148	6	3.4	90	22	AAW71892	Human bone marrow	221	6	3.4	155	22	ABG10871	Novel human diagno
149	6	3.4	90	22	AAE32172	Peptide #6209 enc	222	6	3.4	156	22	ABG04613	Novel human diagno
150	6	3.4	90	22	AAE55895	Skin cell protein,	223	6	3.4	159	22	ABG20837	Novel human diagno
151	6	3.4	90	23	ABG41705	Human peptide enc	224	6	3.4	161	22	ABG12607	Novel human diagno
152	6	3.4	90	23	ABE72095	Rat protein isolat	225	6	3.4	161	22	AAO09519	Human polypeptide
153	6	3.4	91	22	AAW57236	Human brain expres	226	6	3.4	162	18	AAW89772	Staphylococcus aur
154	6	3.4	91	22	AAW29971	Peptide #4008 enc	227	6	3.4	162	12	AAO8128	A polyphenol oxida
155	6	3.4	94	21	AAE16277	Arabidopsis thaliana	228	6	3.4	168	19	AAW75107	Human secreted pro
156	6	3.4	95	21	AAE34553	Human secreted pro	229	6	3.4	170	21	AAO09728	Arabidopsis thaliana

230	6	3.4	174	21	AAB41911	Human ORFX ORF1675	303	6	3.4	267	23	AAO15079	Human GABARib tra
231	6	3.4	175	15	AAR62783	Borrelia VS461 ant	304	6	3.4	268	23	ABP29308	Streptococcus poly
232	6	3.4	175	15	AAR62788	Borrelia 20047 ant	305	6	3.4	269	19	AAV85990	S. pneumoniae deri
233	6	3.4	178	15	AAR62776	Borrelia SIMON ant	306	6	3.4	270	22	ABB69971	Drosophila melanog
234	6	3.4	178	15	AAR60890	Borrelia SIMON ant	307	6	3.4	273	21	AAG36285	Arabidopsis thalia
235	6	3.4	181	23	ABB90622	Chlamydia pneumoni	308	6	3.4	273	21	AAG40304	Arabidopsis thalia
236	6	3.4	181	23	ABB90677	Chlamydia pneumoni	309	6	3.4	276	19	AAW61204	Streptococcus sp07
237	6	3.4	184	21	AAB24445	Human secreted pro	310	6	3.4	276	23	ABP54623	S. pneumoniae pnu
238	6	3.4	185	20	AAV35791	Amino acid sequenc	311	6	3.4	277	21	AAB13593	Streptomyces globi
239	6	3.4	189	15	AAR60909	Borrelia H13 anti9	312	6	3.4	277	23	ABB89958	Human polypeptide
240	6	3.4	191	15	AAR60897	Borrelia VS461 ant	313	6	3.4	278	12	AAH13338	Mac-2 protein inci
241	6	3.4	191	15	AAR60902	Borrelia 20047 ant	314	6	3.4	278	21	AAH56745	Human prostate can
242	6	3.4	191	22	AAU87130	Novel central nerv	315	6	3.4	278	23	AAE23547	Human FAII extrace
243	6	3.4	191	22	AAU86766	Novel human connec	316	6	3.4	280	22	ABB71298	Drosophila melanog
244	6	3.4	191	22	ABP59045	Drosophila melanog	317	6	3.4	282	20	AAV14483	Fragment of human
245	6	3.4	192	22	ABG26443	Novel human diagno	318	6	3.4	288	21	AAG57319	Arabidopsis thalia
246	6	3.4	192	23	ABP00179	Human OREX protein	319	6	3.4	288	23	ABP30012	Streptococcus poly
247	6	3.4	194	22	AAU69472	Human purified sec	320	6	3.4	292	23	ABB79128	Human arginine met
248	6	3.4	196	22	ABG11745	Novel human diagno	321	6	3.4	296	17	AAW11582	Streptomyces prist
249	6	3.4	207	22	ABG20139	Novel human diagno	322	6	3.4	296	21	AAG13941	Arabidopsis thalia
250	6	3.4	214	20	AAV06303	Mouse pancreatic-d	323	6	3.4	296	21	AAG37305	Arabidopsis thalia
251	6	3.4	216	10	AAV93703	Sequence of the 20	324	6	3.4	296	23	ABP92656	Herbicidally activ
252	6	3.4	216	18	AAW31582	Elmeria tenella sp	325	6	3.4	298	21	AAG09727	Arabidopsis thalia
253	6	3.4	217	22	ABB10356	Human CDNA SEQ ID	326	6	3.4	299	16	AAH66405	Bg/1 modification
254	6	3.4	220	21	AAB33001	Pinus radiata tran	327	6	3.4	299	21	AAG09726	Arabidopsis thalia
255	6	3.4	222	22	ABG16504	Novel human diagno	328	6	3.4	302	21	AAG04966	Arabidopsis thalia
256	6	3.4	223	20	AAV36346	Fragment of human	329	6	3.4	304	21	AAG37304	Arabidopsis thalia
257	6	3.4	223	22	ABB12247	Human PI 5-phospha	330	6	3.4	305	21	AAH58916	Breast and ovarian
258	6	3.4	226	22	AAH39182	Human polypeptide	331	6	3.4	305	21	AAG13940	Arabidopsis thalia
259	6	3.4	228	18	AAW20224	H. pylori transpor	332	6	3.4	305	23	AAE23557	Human FAII protein
260	6	3.4	230	22	ABH70669	Drosophila melanog	333	6	3.4	307	22	ABB51275	Human secreted pro
261	6	3.4	231	21	AAG13942	Arabidopsis thalia	334	6	3.4	307	22	ABB51278	Human secreted pro
262	6	3.4	231	21	AAG37306	Arabidopsis thalia	335	6	3.4	308	21	AAG21127	Arabidopsis thalia
263	6	3.4	233	22	AAG93251	Human protein HP02	336	6	3.4	310	23	ABP43485	Human secreted pro
264	6	3.4	234	19	AAW37089	Lycopodium escul	337	6	3.4	311	21	AAH46265	Arabidopsis thalia
265	6	3.4	235	21	AAG29466	Arabidopsis thalia	338	6	3.4	314	20	AAV35552	Chlamydia pneumoni
266	6	3.4	236	22	ABB58682	Drosophila melanog	339	6	3.4	315	22	AAH39424	Human polypeptide
267	6	3.4	237	10	AAV90117	Group B Immunogen	340	6	3.4	318	21	AAG36284	Arabidopsis thalia
268	6	3.4	237	10	AAV90502	Gp.B Elmeria tenel	341	6	3.4	318	21	AAG40303	Arabidopsis thalia
269	6	3.4	237	19	AAV72600	Group B Elmeria te	342	6	3.4	321	22	ABB51277	Human secreted pro
270	6	3.4	239	23	ABP29167	Streptococcus poly	343	6	3.4	321	23	ABG60935	Novel floral meris
271	6	3.4	240	18	AAW25784	Anti-gp130 antibod	344	6	3.4	323	21	AAG07631	Arabidopsis thalia
272	6	3.4	240	18	AAW20555	Helicobacter pylor	345	6	3.4	323	21	AAG18224	Arabidopsis thalia
273	6	3.4	240	18	AAW24687	H. pylori cytoplas	346	6	3.4	323	21	AAG36807	Arabidopsis thalia
274	6	3.4	241	18	AAW20668	H. pylori transpor	347	6	3.4	323	23	ABP93042	Herbicidally activ
275	6	3.4	243	22	ABG13934	Novel human diagno	348	6	3.4	324	21	AAG07630	Arabidopsis thalia
276	6	3.4	247	21	AAH12846	Human V-type ATPas	349	6	3.4	324	21	ABG36806	Arabidopsis thalia
277	6	3.4	247	22	AAU32458	Novel human secret	350	6	3.4	324	22	ABG07010	Novel human diagno
278	6	3.4	248	11	AAH05221	Antigen GX3282(Ex	351	6	3.4	325	21	AAH19393	Amino acid sequenc
279	6	3.4	250	21	AAH42488	Human ORFX ORF2252	352	6	3.4	325	21	AAH05322	Arabidopsis thalia
280	6	3.4	250	21	AAG18225	Arabidopsis thalia	353	6	3.4	325	21	AAG06659	Arabidopsis thalia
281	6	3.4	252	22	AAH40968	Human polypeptide	354	6	3.4	325	21	AAG08960	Arabidopsis thalia
282	6	3.4	252	22	AAH79289	Human protein SEQ	355	6	3.4	325	21	AAG49308	Arabidopsis thalia
283	6	3.4	253	21	AAG36286	Arabidopsis thalia	356	6	3.4	325	21	AAG57818	Arabidopsis thalia
284	6	3.4	253	21	AAG40305	Arabidopsis thalia	357	6	3.4	325	22	AAU71864	C. glutamicum cyst
285	6	3.4	253	22	ABB64223	Drosophila melanog	358	6	3.4	325	22	AAU72281	C. glutamicum proce
286	6	3.4	254	21	AAG21129	Arabidopsis thalia	359	6	3.4	326	21	AAH23045	Human peroxisome a
287	6	3.4	258	19	AAW61954	Rat galectin amino	360	6	3.4	326	21	AAG46264	Arabidopsis thalia
288	6	3.4	259	20	AAV34948	Chlamydia pneumoni	361	6	3.4	326	22	AAH41210	Human polypeptide
289	6	3.4	259	21	AAG21128	Arabidopsis thalia	362	6	3.4	327	21	AAH18223	Arabidopsis thalia
290	6	3.4	261	22	ABP60232	Drosophila melanog	363	6	3.4	327	21	AAG57318	Arabidopsis thalia
291	6	3.4	262	14	AAH42200	IgE binding protei	364	6	3.4	328	20	AAV74117	Human prostate tum
292	6	3.4	262	22	AAU25441	Human mddt protein	365	6	3.4	329	21	AAG44530	Arabidopsis thalia
293	6	3.4	263	23	AAU97821	Rat cell membrane	366	6	3.4	329	23	AAH50863	Arabidopsis thalia
294	6	3.4	263	23	AAU97821	Mouse cell membran	367	6	3.4	330	21	AAG07629	Arabidopsis thalia
295	6	3.4	264	12	AAH12531	Mac2.16 expression	368	6	3.4	330	21	AAG36805	Arabidopsis thalia
296	6	3.4	264	19	AAW71219	Beta-D-galactoside	369	6	3.4	330	22	AAU57276	Propionibacterium
297	6	3.4	265	22	ABH88462	Human membrane or	370	6	3.4	332	22	AAH49234	Mouse SNORF33 rece
298	6	3.4	267	15	AAH56259	K. marxianus URA3.	371	6	3.4	333	21	AAV44632	A. thaliana putati
299	6	3.4	267	20	AAV49111	Human GABARib tra	372	6	3.4	333	22	AAH85457	Arabidopsis thalia
300	6	3.4	267	20	AAV49112	Human GABARib tra	373	6	3.4	336	21	AAG46263	Arabidopsis thalia
301	6	3.4	267	22	AAH93536	Human polypeptide,	374	6	3.4	340	21	AAG04965	Arabidopsis thalia
302	6	3.4	267	23	AAO15078	Human GABARib tra	375	6	3.4	340	21	AAG08959	Arabidopsis thalia

376	6	3.4	340	21	AA57817	Arabidopsis thalia	449	6	3.4	407	22	ABG21343	Novel human diagno
377	6	3.4	340	22	AA01012	CPE II protein seq	450	6	3.4	409	22	ABG10439	Novel human diagno
378	6	3.4	340	23	AAE22586	BS203 consensus pr	451	6	3.4	410	20	AA537660	Chlamydia trachoma
379	6	3.4	341	19	AAW58573	Trichoderma reesel	452	6	3.4	410	22	AB04102	Botulism toxin hea
380	6	3.4	341	21	AA57816	Arabidopsis thalia	453	6	3.4	411	21	AA576047	Rat skin cell prot
381	6	3.4	342	23	AAE22916	Human transporter	454	6	3.4	411	22	AA55986	Skin cell protein,
382	6	3.4	342	23	AAE23546	Human mature FAIL	455	6	3.4	411	23	AB72186	Rat protein isolat
383	6	3.4	346	21	AA36678	Arabidopsis thalia	456	6	3.4	413	21	AA04964	Arabidopsis thalia
384	6	3.4	346	23	AB91294	Herbicidally activ	457	6	3.4	414	21	AA515375	Arabidopsis thalia
385	6	3.4	349	22	ABG16891	Novel human diagno	458	6	3.4	415	22	AA40868	Human polypeptide
386	6	3.4	349	22	AA67525	Amino acid sequenc	459	6	3.4	416	20	AA88745	Secreted protein e
387	6	3.4	350	21	AA36677	Arabidopsis thalia	460	6	3.4	416	22	AB50513	Human secreted pro
388	6	3.4	350	23	AB53286	Human polypeptide	461	6	3.4	416	22	AA25861	Human protein sequ
389	6	3.4	352	21	AA05321	Arabidopsis thalia	462	6	3.4	417	21	AA41474	Human OREX ORF1238
390	6	3.4	352	21	AA049307	Arabidopsis thalia	463	6	3.4	417	21	AA41474	Human membrane or
391	6	3.4	353	16	AA75388	Japanese cedar pol	464	6	3.4	418	22	AA88368	Botulism toxin hea
392	6	3.4	353	17	AA81587	Cedar pollen aller	465	6	3.4	419	22	AA04095	Botulism toxin hea
393	6	3.4	358	21	AA94336	Human cell surface	466	6	3.4	421	22	AA63222	Amino acid sequenc
394	6	3.4	359	20	AA41690	Human PRO329 prote	467	6	3.4	426	23	AAE23289	Human nectin-3alph
395	6	3.4	359	21	AA44246	Human PRO329 (UNQ2	468	6	3.4	432	21	AAE23289	Arabidopsis thalia
396	6	3.4	359	21	AA34744	Human secreted pro	469	6	3.4	432	21	AAE15374	Arabidopsis thalia
397	6	3.4	359	22	AAU29044	Human PRO polypept	470	6	3.4	433	21	AAU36234	Arabidopsis thalia
398	6	3.4	359	22	AA03451	Human gene 25 enco	471	6	3.4	434	22	AAU36234	Pseudomonas aerugi
399	6	3.4	359	23	ABG3386	Human albumin fusi	472	6	3.4	435	21	AAE14364	Arabidopsis thalia
400	6	3.4	359	23	ABP39997	Staphylococcus epi	473	6	3.4	435	22	AB11390	Human secreted pro
401	6	3.4	362	21	AA06658	Arabidopsis thalia	474	6	3.4	435	22	AA80099	Human protein SEQ
402	6	3.4	362	21	AA36676	Arabidopsis thalia	475	6	3.4	436	22	ABG18343	Novel human diagno
403	6	3.4	363	20	AA537863	Amino acid sequenc	476	6	3.4	437	23	AAE23299	Human nectin-3gamm
404	6	3.4	363	20	AA53633	Human secreted pro	477	6	3.4	438	22	AAE33984	Amino acid sequenc
405	6	3.4	363	20	AA88563	Secreted protein e	478	6	3.4	438	23	AAE23293	Mouse nectin-3gamm
406	6	3.4	363	22	AB50330	Human secreted pro	479	6	3.4	442	22	AA82974	S. epidermidis ope
407	6	3.4	363	22	AB51276	Human secreted pro	480	6	3.4	443	22	AAE00483	Streptomyces galli
408	6	3.4	364	21	AA329465	Arabidopsis thalia	481	6	3.4	447	21	AAE15496	Arabidopsis thalia
409	6	3.4	364	22	ABG14391	Novel human diagno	482	6	3.4	448	23	AA894438	Human protein sequ
410	6	3.4	365	21	AA06657	Arabidopsis thalia	483	6	3.4	448	23	ABG61513	Iron uptake ABC tr
411	6	3.4	365	21	AA08958	Arabidopsis thalia	484	6	3.4	449	21	AAE14363	Arabidopsis thalia
412	6	3.4	370	20	AAW88777	Polypeptide fragme	485	6	3.4	449	21	AA577137	Synthetic botulinu
413	6	3.4	370	22	AB50610	Human secreted pro	486	6	3.4	450	22	AA804094	Botulism toxin hea
414	6	3.4	370	22	AAE23544	Human FAIL protein	487	6	3.4	450	22	ABG64406	Drosophila melanog
415	6	3.4	370	23	AAE23553	Human FAIL protein	488	6	3.4	451	19	AAW68395	Clostridium botuli
416	6	3.4	370	23	AAE23554	Human FAIL protein	489	6	3.4	452	19	AAW68395	Clostridium botuli
417	6	3.4	370	23	AAE23555	Human FAIL protein	490	6	3.4	452	21	AA811122	Clostridium botuli
418	6	3.4	370	23	AAE23556	Human FAIL protein	491	6	3.4	452	21	AA575750	Human growth-assoc
419	6	3.4	372	21	AAE57317	Arabidopsis thalia	492	6	3.4	455	19	AAW76439	Human p53 regulat
420	6	3.4	373	20	AA525664	Cedar allergen 493	493	6	3.4	456	23	AB808386	Tmox amino acid se
421	6	3.4	373	20	AA525668	Japanese cedar all	494	6	3.4	458	22	AAU29271	Human PRO polypept
422	6	3.4	374	14	AA831937	Cry j I. Cryptome	495	6	3.4	459	12	AA810099	Beta subunit of AT
423	6	3.4	374	15	AA845541	Cry j I pollen all	496	6	3.4	465	21	AA831991	Arabidopsis thalia
424	6	3.4	374	15	AA860166	Japanese cedar pol	497	6	3.4	465	22	AAU68517	Human novel cytoke
425	6	3.4	374	16	AA82490	Cry j I Japanese C	498	6	3.4	465	22	AA82790	Human protein sequ
426	6	3.4	374	20	AA525665	Cedar allergen 493	499	6	3.4	466	22	AA895611	Human protein sequ
427	6	3.4	374	20	AA525669	Japanese cedar all	500	6	3.4	468	13	AA825597	PHO. Pichia pasto
428	6	3.4	379	21	AA849306	Arabidopsis thalia	501	6	3.4	469	22	AB88765	Drosophila melanog
429	6	3.4	379	23	AB891647	Herbicidally activ	502	6	3.4	476	20	AA535603	Chlamydia pneumoni
430	6	3.4	380	21	AA053320	Arabidopsis thalia	503	6	3.4	476	21	AA836464	Arabidopsis thalia
431	6	3.4	382	23	ABP25521	Streptococcus poly	504	6	3.4	476	22	AAU38890	C. pneumoniae CT79
432	6	3.4	383	23	AB854588	Lactococcus lactis	505	6	3.4	477	20	AAU37784	Protein involved i
433	6	3.4	386	21	AA444631	A. thaliana AIR sy	506	6	3.4	481	21	AA588312	C. parvum protein
434	6	3.4	386	22	AA854556	Arabidopsis AIR sy	507	6	3.4	483	22	AA000982	Human bone marrow
435	6	3.4	387	23	AAE23290	Human nectin-3beta	508	6	3.4	485	21	AA836463	Arabidopsis thalia
436	6	3.4	393	20	AA536194	Human secreted pro	509	6	3.4	486	22	AA579736	Human protein SEQ
437	6	3.4	393	21	AA594883	Human protein clon	510	6	3.4	490	20	AA533929	Leukotoxin carrier
438	6	3.4	393	22	AAW79115	Human protein sequ	511	6	3.4	490	22	AB844563	Human wound healin
439	6	3.4	393	22	AA893277	Human protein sequ	512	6	3.4	490	22	AAU79767	Human protein SEQ
440	6	3.4	393	22	AA861136	Human NOV6 protei	513	6	3.4	490	23	AB53621	Herbicidally activ
441	6	3.4	394	22	AA890342	C glutamicum prote	514	6	3.4	494	20	AA534615	Chlamydia pneumoni
442	6	3.4	394	22	AA895622	Human protein sequ	515	6	3.4	494	22	ABG14289	Novel human diagno
443	6	3.4	394	23	AB892846	Herbicidally activ	516	6	3.4	494	23	AB890592	Chlamydia pneumoni
444	6	3.4	395	22	AAU38229	Salmonella typhi c	517	6	3.4	494	23	AB853680	Lactococcus lactis
445	6	3.4	401	22	AAU38235	Salmonella typhi c	518	6	3.4	495	20	AA533296	Human membrane spa
446	6	3.4	404	20	AA888317	Sugar transferase	519	6	3.4	495	23	AB853285	Human polypeptide
447	6	3.4	405	21	AA842465	Human OREX ORF2229	520	6	3.4	496	18	AAU33901	Streptococcus pneu
448	6	3.4	406	22	ABG00659	Novel human diagno	521	6	3.4	496	22	AAU37808	Streptococcus pneu
													Amino acid sequenc

522	6	3.4	497	22	AAG82265	S. epidermidis ope	595	6	3.4	560	23	AAU77628	S. agalactiae Sbp2
523	6	3.4	498	22	ABG23519	Novel human diagno	596	6	3.4	563	20	AAV14111	Human GABAB recept
524	6	3.4	501	23	ABP25949	Streptococcus poly	597	6	3.4	563	22	ABY72358	Virulent group B S
525	6	3.4	502	22	AAU33945	Staphylococcus aur	598	6	3.4	565	22	ABG18709	Novel human diagno
526	6	3.4	502	22	AAU36794	Staphylococcus aur	599	6	3.4	569	23	ABB92340	Herbicidially activ
527	6	3.4	502	22	AAU37228	Staphylococcus aur	600	6	3.4	571	21	AAG31041	Arabidopsis thalia
528	6	3.4	502	22	AAW79308	Human protein SEQ	601	6	3.4	574	23	ABP97289	Novel human protei
529	6	3.4	503	22	AAG82358	S. epidermidis ope	602	6	3.4	575	15	AAW54220	L.lactis alpha-ace
530	6	3.4	503	23	ABP33639	Staphylococcus epi	603	6	3.4	577	22	AAW78367	Human protein SEQ
531	6	3.4	504	22	ABW52616	Escherichia coli p	604	6	3.4	579	22	AAE06804	Mature human neuro
532	6	3.4	504	23	AAE23284	Human deleted nect	605	6	3.4	580	22	ABG18705	Novel human diagno
533	6	3.4	507	21	AAE36462	Arabidopsis thalia	606	6	3.4	581	23	ABB93540	Herbicidially activ
534	6	3.4	507	21	AAW81580	Streptococcus pneu	607	6	3.4	582	22	ABG06711	Novel human diagno
535	6	3.4	507	23	ABP40226	Staphylococcus epi	608	6	3.4	583	22	ABG16197	Novel human diagno
536	6	3.4	508	21	AAE14362	Arabidopsis thalia	609	6	3.4	584	22	AAW00930	Human bone marrow
537	6	3.4	510	22	AAG63983	Amino acid sequenc	610	6	3.4	585	22	AAW79351	Human protein SEQ
538	6	3.4	510	23	AAE23285	Mouse nectin-3-hum	611	6	3.4	586	22	ABG13054	Novel human diagno
539	6	3.4	510	23	AAE23286	Human nectin-3beta	612	6	3.4	588	17	AAW05297	Esterase secretory
540	6	3.4	510	23	AAE23292	Mouse nectin-3beta	613	6	3.4	594	19	AAW74580	Human membrane pro
541	6	3.4	510	23	AAU75483	S. aureus antigeni	614	6	3.4	594	22	ABG26986	Novel human diagno
542	6	3.4	511	21	AAG58809	Arabidopsis thalia	615	6	3.4	594	22	ABG94068	Human protein sequ
543	6	3.4	512	19	AAW68506	Human acid sensinq	616	6	3.4	595	21	AAG58807	Arabidopsis thalia
544	6	3.4	512	20	AAW93420	Human BNC1 protein	617	6	3.4	595	23	AAE23288	Human nectin-3beta
545	6	3.4	512	21	AAW69177	A human acid-sensi	618	6	3.4	597	21	AAW08122	A polyphenol oxida
546	6	3.4	512	23	AAU10903	Human acid sensinq	619	6	3.4	606	22	AAE06799	Human neuronal gui
547	6	3.4	512	23	AAU10905	Rat acid sensinq i	620	6	3.4	606	23	ABP60969	Novel human protei
548	6	3.4	515	22	AAW80287	Human protein SEQ	621	6	3.4	606	23	ABP60970	Novel human protei
549	6	3.4	516	22	ABG18191	Novel human diagno	622	6	3.4	606	23	AAU96849	Human NOV4a protei
550	6	3.4	517	23	ABP41923	Human ovarian anti	623	6	3.4	606	23	AAU96850	Human NOV4b protei
551	6	3.4	518	23	ABB54962	Lactococcus lactis	624	6	3.4	606	23	ABG60229	Human NOV4b protei
552	6	3.4	520	13	AAK21075	Sequence encoded b	625	6	3.4	606	23	ABG60230	Human Silt-3-like
553	6	3.4	520	13	AAK21076	Sequence encoded b	626	6	3.4	606	23	AAU85410	Human Silt-3-like
554	6	3.4	520	13	AAK21077	Sequence encoded b	627	6	3.4	606	23	AAU85411	Human protein NOV8
555	6	3.4	521	21	AAE13578	Streptomyces globi	628	6	3.4	606	23	AAU85418	Human protein NOV9
556	6	3.4	521	21	AAE13605	Streptomyces globi	629	6	3.4	611	22	AAE85021	NOV8 allelic varia
557	6	3.4	521	21	AAE13605	Streptomyces globi	630	6	3.4	612	21	AAE85021	Shrimp white spot
558	6	3.4	524	21	AAE13605	Human ORFX ORF164	631	6	3.4	612	21	AAE85021	Arabidopsis thalia
559	6	3.4	524	21	AAE13605	Human transmembran	632	6	3.4	613	22	AAW02556	Human polypeptide
560	6	3.4	524	23	ABP39470	Staphylococcus epi	633	6	3.4	614	19	AAW38320	Specific co-activa
561	6	3.4	529	21	AAE22200	Human receptor mol	634	6	3.4	614	20	AAW84596	Amino acid sequenc
562	6	3.4	533	19	AAW59442	Hordeum vulgare ML	635	6	3.4	620	20	AAV13357	Amino acid sequenc
563	6	3.4	533	19	AAW59443	Hordeum vulgare ML	636	6	3.4	620	20	AAU12333	Human PRO227 poly
564	6	3.4	533	21	AAW03401	Barley Mlo protein	637	6	3.4	620	22	AAU12333	Human membrane ass
565	6	3.4	534	20	AAE26966	Wheat Mlo fungal r	638	6	3.4	621	19	AAW62842	Human PRO227 prote
566	6	3.4	534	20	AAE26968	Wheat Mlo fungal r	639	6	3.4	621	21	AAE11121	Helicobacter pylor
567	6	3.4	534	21	AAE01805	Wheat Mlo homologi	640	6	3.4	621	21	AAE11121	Human GAMT-1 prote
568	6	3.4	534	22	AAE01805	Amino acid sequenc	641	6	3.4	621	21	AAE11121	Human growth-assoc
569	6	3.4	534	22	AAE01805	Amino acid sequenc	642	6	3.4	621	22	AAU35710	Helicobacter pylor
570	6	3.4	534	22	AAE01805	Amino acid sequenc	643	6	3.4	621	22	AAU35710	Human protein sequ
571	6	3.4	537	23	ABP28881	Streptococcus poly	644	6	3.4	622	10	AAE91632	Rhoptry membrane a
572	6	3.4	537	23	ABP28881	Streptococcus poly	645	6	3.4	622	13	AAE27532	Plasmodium falcipa
573	6	3.4	540	23	ABB54556	Lactococcus lactis	646	6	3.4	622	16	AAE68840	Thermococcus amida
574	6	3.4	542	23	AAE23281	Human deleted nect	647	6	3.4	622	19	AAW41248	Human polypeptide,
575	6	3.4	544	17	AAW03343	LKT-GnRH protein f	648	6	3.4	622	22	AAW93427	Human polypeptide,
576	6	3.4	544	19	AAW79570	LKT-GnRH chimeric	649	6	3.4	622	22	AAW93427	Human polypeptide,
577	6	3.4	545	22	AAE27253	Streptococcus poly	650	6	3.4	622	23	AAU99359	Synthetic P. falcic
578	6	3.4	546	21	AAE27253	Human polypeptide	651	6	3.4	622	23	AAU99359	P. falciparum apic
579	6	3.4	546	21	AAE27253	A. pleuropneumonia	652	6	3.4	622	23	ABP43477	Human secreted pro
580	6	3.4	547	14	AAE27253	Ap Serotype 7 60kD	653	6	3.4	622	21	AAE27253	Arabidopsis thalia
581	6	3.4	547	22	AAE27253	Human AFP protein	654	6	3.4	626	22	AAE27253	Human protein sequ
582	6	3.4	548	22	AAE27253	Amino acid sequenc	655	6	3.4	627	21	AAE27253	Arabidopsis thalia
583	6	3.4	549	22	AAE27253	Amino acid sequenc	656	6	3.4	634	23	AAE23287	Human nectin-3alp
584	6	3.4	549	22	AAE27253	Amino acid sequenc	657	6	3.4	638	21	AAE23287	Human nectin-3alp
585	6	3.4	549	23	AAE27253	Mouse nectin-3-hum	658	6	3.4	642	21	AAE23287	Nelisseria gonorrhe
586	6	3.4	549	23	AAE27253	Human nectin-3alp	659	6	3.4	644	20	AAE23287	Human ORFX ORF2886
587	6	3.4	549	23	AAE27253	Mouse nectin-3alp	660	6	3.4	650	21	AAE23287	C. pneumoniae prot
588	6	3.4	550	21	AAE27253	Mouse nectin-3alp	661	6	3.4	650	21	AAE23287	Arabidopsis thalia
589	6	3.4	550	22	AAE27253	Arabidopsis thalia	662	6	3.4	652	21	AAE23287	Human RPI05 protei
590	6	3.4	550	22	AAE27253	Human polypeptide	663	6	3.4	654	22	AAE23287	Arabidopsis thalia
591	6	3.4	550	22	AAE27253	Drosophila melanog	664	6	3.4	654	22	AAE23287	Human polypeptide
592	6	3.4	550	22	AAE27253	Human polypeptide	665	6	3.4	658	22	AAE23287	Enterococcus faeca
593	6	3.4	559	22	AAE27253	Human polypeptide	666	6	3.4	661	18	AAW28510	Product of clone J
594	6	3.4	559	22	AAE27253	Human secreted pro	667	6	3.4	661	19	AAW47274	Human B-cell activ
			560	21	AAE27253	Human ORFX ORF2802				662	20	AAW75556	B cell surface prot
										662	16	AAW75556	Cotransporter prot

668	6	3.4	664	22	AAU36984	Staphylococcus aur	741	6	3.4	843	22	ABG10140	Novel human diagno
669	6	3.4	668	23	ABB91145	Herbicidally activ	742	6	3.4	844	19	AAW40118	Rat GABA-BR1b rece
670	6	3.4	671	20	AAV35050	Chlamydia pneumoni	743	6	3.4	844	19	AAW40119	Human GABA-BR1b re
671	6	3.4	676	23	ABB47730	Listeria monocytog	744	6	3.4	844	20	AAV28839	Human GABABR1b rec
672	6	3.4	678	21	AGC30708	Arabidopsis thalia	745	6	3.4	844	20	AAV28842	Rat GABABR1b recep
673	6	3.4	678	21	AGC30708	Arabidopsis thalia	746	6	3.4	844	20	AAV49123	Rat GABABR1b prote
674	6	3.4	684	16	AAR73032	Human Factor-VIII	747	6	3.4	844	20	AAV49123	Human GABAB recep
675	6	3.4	684	16	AAR73032	Human Factor-VIII	748	6	3.4	844	20	AAV49123	Human G-protein co
676	6	3.4	685	13	AAV25911	Macrophage tissue	749	6	3.4	844	23	AAV32467	Rat GABABR1b prote
677	6	3.4	685	13	AAV25911	Macrophage tissue	750	6	3.4	844	23	AAO15090	Rat GABABR1b recep
678	6	3.4	685	23	ABB57342	Murine tissue tran	751	6	3.4	844	23	AAO14255	Human GABA-B recep
679	6	3.4	691	11	AAR04711	Mouse ischaemic co	752	6	3.4	844	23	ABB05634	Human GABA-B recep
680	6	3.4	691	20	AAV42329	Sequence of guinea	753	6	3.4	848	23	ABG69087	Botulinum neurotox
681	6	3.4	691	20	AAV42329	Staphylococcus aur	754	6	3.4	859	21	AAV30786	Arabidopsis thalia
682	6	3.4	691	22	AAU33182	Staphylococcus aur	755	6	3.4	866	21	AAV30786	Arabidopsis thalia
683	6	3.4	691	22	AAU33182	Staphylococcus aur	756	6	3.4	866	21	AAV30786	Arabidopsis thalia
684	6	3.4	691	22	AAU33182	Staphylococcus aur	757	6	3.4	866	21	AAV30786	Arabidopsis thalia
685	6	3.4	692	22	AAU33182	Human clathrin 76	758	6	3.4	868	22	ABB68310	Neisseria gonorrhoe
686	6	3.4	692	22	AAU33182	Human protein sequ	759	6	3.4	868	22	ABB68310	Drosophila melanog
687	6	3.4	693	20	AAV30797	A mouse transgluta	760	6	3.4	871	23	ABG68285	Brevibacillus late
688	6	3.4	693	20	AAV30797	A human transgluta	761	6	3.4	872	23	ABG68285	AnEPV mRNA capping
689	6	3.4	695	19	AAV79573	LKT-GnRH chimeric	762	6	3.4	882	22	AAV30786	A human alpha-2 ma
690	6	3.4	695	21	AAV58133	Leukotoxin/gonadot	763	6	3.4	882	22	AAV30786	Sequence of low mo
691	6	3.4	695	21	AAV58133	Gonadotropin relea	764	6	3.4	884	14	AAV34402	Human GABAB recep
692	6	3.4	697	21	AAV93320	Amino acid sequenc	765	6	3.4	886	20	ABV71249	Drosophila melanog
693	6	3.4	697	21	AAV93320	Amino acid sequenc	766	6	3.4	890	22	ABV71249	Human GABAB recep
694	6	3.4	697	21	AAV93320	Novel human diagno	767	6	3.4	892	20	AAV14109	Human GABAB recep
695	6	3.4	701	22	ABG13634	Novel human diagno	768	6	3.4	892	20	AAV14109	Bacteriophage RB69
696	6	3.4	701	22	ABG13634	Novel human diagno	769	6	3.4	892	20	AAV14109	Human GABAB recep
697	6	3.4	701	22	ABG13634	Novel human diagno	770	6	3.4	892	20	AAV14109	A human alpha-2 ma
698	6	3.4	701	22	ABG13634	Novel human diagno	771	6	3.4	892	20	AAV14109	Human MERCK hgbic
699	6	3.4	701	22	ABG13634	Novel human diagno	772	6	3.4	892	20	AAV14109	Human GABAB recep
700	6	3.4	701	22	ABG13634	Novel human diagno	773	6	3.4	892	20	AAV14109	Human GABAB recep
701	6	3.4	701	22	ABG13634	Novel human diagno	774	6	3.4	892	20	AAV14109	Novel human diagno
702	6	3.4	701	22	ABG13634	Novel human diagno	775	6	3.4	892	20	AAV14109	N. meningitidis str
703	6	3.4	701	22	ABG13634	Novel human diagno	776	6	3.4	892	20	AAV14109	N. meningitidis st
704	6	3.4	701	22	ABG13634	Novel human diagno	777	6	3.4	892	20	AAV14109	N. meningitidis tra
705	6	3.4	701	22	ABG13634	Novel human diagno	778	6	3.4	892	20	AAV14109	Neisseria meningit
706	6	3.4	701	22	ABG13634	Novel human diagno	779	6	3.4	892	20	AAV14109	A human alpha-2 ma
707	6	3.4	701	22	ABG13634	Novel human diagno	780	6	3.4	892	20	AAV14109	Novel human diagno
708	6	3.4	701	22	ABG13634	Novel human diagno	781	6	3.4	892	20	AAV14109	Phosphoenolpyruvat
709	6	3.4	701	22	ABG13634	Novel human diagno	782	6	3.4	892	20	AAV14109	Phosphoenolpyruvat
710	6	3.4	701	22	ABG13634	Novel human diagno	783	6	3.4	892	20	AAV14109	C glutamicum prote
711	6	3.4	701	22	ABG13634	Novel human diagno	784	6	3.4	892	20	AAV14109	Arabidopsis thalia
712	6	3.4	701	22	ABG13634	Novel human diagno	785	6	3.4	892	20	AAV14109	Leukotoxin 352 enc
713	6	3.4	701	22	ABG13634	Novel human diagno	786	6	3.4	892	20	AAV14109	Leukotoxin 352 pro
714	6	3.4	701	22	ABG13634	Novel human diagno	787	6	3.4	892	20	AAV14109	Recombinant leukot
715	6	3.4	701	22	ABG13634	Novel human diagno	788	6	3.4	892	20	AAV14109	Recombinant leukot
716	6	3.4	701	22	ABG13634	Novel human diagno	789	6	3.4	892	20	AAV14109	Recombinant leukot
717	6	3.4	701	22	ABG13634	Novel human diagno	790	6	3.4	892	20	AAV14109	Corynebacterium gl
718	6	3.4	701	22	ABG13634	Novel human diagno	791	6	3.4	892	20	AAV14109	Chlamydia pneumoni
719	6	3.4	701	22	ABG13634	Novel human diagno	792	6	3.4	892	20	AAV14109	LKT352. Pasteurel
720	6	3.4	701	22	ABG13634	Novel human diagno	793	6	3.4	892	20	AAV14109	Leukotoxin 352 pro
721	6	3.4	701	22	ABG13634	Novel human diagno	794	6	3.4	892	20	AAV14109	Recombinant leukot
722	6	3.4	701	22	ABG13634	Novel human diagno	795	6	3.4	892	20	AAV14109	Recombinant leukot
723	6	3.4	701	22	ABG13634	Novel human diagno	796	6	3.4	892	20	AAV14109	P. haemolytica tru
724	6	3.4	701	22	ABG13634	Novel human diagno	797	6	3.4	892	20	AAV14109	Leukotoxin 352 pol
725	6	3.4	701	22	ABG13634	Novel human diagno	798	6	3.4	892	20	AAV14109	Human KIAAN ligase
726	6	3.4	701	22	ABG13634	Novel human diagno	799	6	3.4	892	20	AAV14109	GnRH-leukotoxin ge
727	6	3.4	701	22	ABG13634	Novel human diagno	800	6	3.4	892	20	AAV14109	Chlamydia pneumoni
728	6	3.4	701	22	ABG13634	Novel human diagno	801	6	3.4	892	20	AAV14109	Chlamydia pneumoni
729	6	3.4	701	22	ABG13634	Novel human diagno	802	6	3.4	892	20	AAV14109	Protein of App rel
730	6	3.4	701	22	ABG13634	Novel human diagno	803	6	3.4	892	20	AAV14109	Arabidopsis thalia
731	6	3.4	701	22	ABG13634	Novel human diagno	804	6	3.4	892	20	AAV14109	Somatostatin-leuko
732	6	3.4	701	22	ABG13634	Novel human diagno	805	6	3.4	892	20	AAV14109	Chlamydia pneumoni
733	6	3.4	701	22	ABG13634	Novel human diagno	806	6	3.4	892	20	AAV14109	Amino acid sequenc
734	6	3.4	701	22	ABG13634	Novel human diagno	807	6	3.4	892	20	AAV14109	Drosophila melanog
735	6	3.4	701	22	ABG13634	Novel human diagno	808	6	3.4	892	20	AAV14109	Rotavirus VP4-leuk
736	6	3.4	701	22	ABG13634	Novel human diagno	809	6	3.4	892	20	AAV14109	105KD PTX protein
737	6	3.4	701	22	ABG13634	Novel human diagno	810	6	3.4	892	20	AAV14109	Leukotoxin from P.
738	6	3.4	701	22	ABG13634	Novel human diagno	811	6	3.4	892	20	AAV14109	Leukotoxin protein
739	6	3.4	701	22	ABG13634	Novel human diagno	812	6	3.4	892	20	AAV14109	PtxA protein of Pa
740	6	3.4	701	22	ABG13634	Novel human diagno	813	6	3.4	892	20	AAV14109	Pasteurella haemol

814	6	3.4	960	20	AAV28841	Rat GABABR1a recep	887	6	3.4	1375	23	AAU79288	Streptococcus muta
815	6	3.4	960	20	AAV49122	Rat GABABR1a prote	888	6	3.4	1382	23	ABE76484	Human prostate-spe
816	6	3.4	960	20	AAV29797	Murine gamma-amino	889	6	3.4	1383	18	AAW33227	Procoagulant-activ
817	6	3.4	960	21	AAV83145	Human GABAB1A rec	890	6	3.4	1383	18	AAW33228	Procoagulant-activ
818	6	3.4	960	22	AAV50089	Murine GABA-B-R1a	891	6	3.4	1383	18	AAW33228	Procoagulant-activ
819	6	3.4	960	23	AAO15089	Murine GABA-B-R1a	892	6	3.4	1403	19	AAW79574	Leukotoxin 101. S
820	6	3.4	960	23	ABO5633	Murine GABA-B rece	893	6	3.4	1424	9	AAO2668	Modified factor VII
821	6	3.4	961	20	AAV28838	Human GABABR1a rec	894	6	3.4	1424	10	AAV91169	Sequence of 740 Ar
822	6	3.4	961	20	AAV29798	Human gamma-amino-	895	6	3.4	1424	22	AAV48842	Mutant mature huma
823	6	3.4	961	20	AAV14101	Human GABAB recep	896	6	3.4	1425	9	AAV80267	Modified factor VII
824	6	3.4	961	23	AAO14254	Human GABAB recep	897	6	3.4	1431	22	AAV67960	Amino acid sequenc
825	6	3.4	961	23	ABO5632	Human GABA-B recep	898	6	3.4	1438	21	AAO1262	B-factor deleted f
826	6	3.4	962	22	AAO50090	Human GABA-B-R1a	899	6	3.4	1440	12	AAV12971	Factor VIII:SQ
827	6	3.4	964	20	AAV14105	Canine GABAB recep	900	6	3.4	1443	19	AAV44137	Homo sapiens facto
828	6	3.4	964	23	AAU72538	Arabidopsis cell c	901	6	3.4	1443	20	AAV31598	Porcine factor VII
829	6	3.4	969	22	ABE63009	Drosophila melanog	902	6	3.4	1443	22	AAV50469	Porcine factor VII
830	6	3.4	975	22	ABE60615	Drosophila melanog	903	6	3.4	1457	19	AAW46246	Human factor VIII
831	6	3.4	975	22	ABE93230	Human protein sequ	904	6	3.4	1457	19	AAW44372	Human factor VIII
832	6	3.4	976	22	ABE63981	Drosophila melanog	905	6	3.4	1457	20	AAV21675	Beta-domain delete
833	6	3.4	977	17	AAW03942	LKT-GnRH protein f	906	6	3.4	1459	22	AAE10827	Human factor VIII
834	6	3.4	977	19	AAW79569	LKT-GnRH chimeric	907	6	3.4	1459	22	AAE10832	Human factor VIII
835	6	3.4	977	22	AAW39082	Human polypeptide	908	6	3.4	1459	22	AAE10833	Human factor VIII
836	6	3.4	992	16	AAV73917	Rubella virus Ther	909	6	3.4	1467	22	AAE11207	Modified porcine f
837	6	3.4	1006	22	AAV81530	S. epidermidis ope	910	6	3.4	1467	22	AAV38825	Neisseria gonorrhoe
838	6	3.4	1014	8	AAV71139	Factor VIII:c varia	911	6	3.4	1470	22	AAV31518	Amino acid sequenc
839	6	3.4	1017	23	AAU79285	Streptococcus muta	912	6	3.4	1471	18	AAW23414	Human B-domain del
840	6	3.4	1018	15	AAV63759	Human contactin (E	913	6	3.4	1471	22	AAV67959	Amino acid sequenc
841	6	3.4	1018	17	AAV87028	Human N-methyl-D-a	914	6	3.4	1475	22	AAE21224	Novel human diagno
842	6	3.4	1061	20	AAV87504	Human N-methyl-D-a	915	6	3.4	1475	23	AAU98027	S. mutans glucosyl
843	6	3.4	1069	15	AAV52748	Bovine IFNgamma/LK	916	6	3.4	1475	23	AAU98030	S. mutans glucosyl
844	6	3.4	1069	18	AAV13867	Chimeric protein #	917	6	3.4	1475	23	AAU98031	S. mutans glucosyl
845	6	3.4	1069	21	AAV21074	Bovine gamma-IFN/P	918	6	3.4	1475	23	AAU98032	S. mutans glucosyl
846	6	3.4	1076	23	ABP39192	Staphylococcus epl	919	6	3.4	1475	23	AAU98033	S. mutans glucosyl
847	6	3.4	1081	21	AAV26240	Human N-methyl-D-a	920	6	3.4	1475	23	AAU98034	S. mutans glucosyl
848	6	3.4	1081	23	ABE56523	Human NMDA recepto	921	6	3.4	1475	23	AAU98035	S. mutans glucosyl
849	6	3.4	1086	9	AAV22507	Pullulanase protei	922	6	3.4	1475	23	AAU98036	S. mutans glucosyl
850	6	3.4	1098	13	AAV22103	Bovine IL-2 - LKT	923	6	3.4	1475	23	AAU98037	S. mutans glucosyl
851	6	3.4	1098	15	AAV52747	Bovine IL-2/LKT ch	924	6	3.4	1475	23	AAU98038	S. mutans GTFB mut
852	6	3.4	1098	18	AAV13866	Chimeric protein #	925	6	3.4	1475	23	AAU98039	S. mutans glucosyl
853	6	3.4	1098	21	AAV21073	Bovine IL-2/Pasteu	926	6	3.4	1475	23	AAU98040	S. mutans glucosyl
854	6	3.4	1112	22	ABE62695	Drosophila melanog	927	6	3.4	1476	23	AAU79284	Streptococcus muta
855	6	3.4	1129	22	ABG20477	Novel human diagno	928	6	3.4	1492	23	ABG61768	Novel alpha-2-macr
856	6	3.4	1147	21	AAV30843	Arabidopsis thalia	929	6	3.4	1496	21	AAV43629	Arabidopsis thalia
857	6	3.4	1164	22	AAU03555	Human protein kina	930	6	3.4	1508	22	AAV63548	A human alpha-2 ma
858	6	3.4	1164	23	AAV19153	Human kinase poly	931	6	3.4	1511	21	AAV46260	Arabidopsis thalia
859	6	3.4	1172	22	ABE66459	Drosophila melanog	932	6	3.4	1514	23	ABV90837	Herbicidially activ
860	6	3.4	1212	20	AAV87503	Human N-methyl-D-a	933	6	3.4	1516	9	AAV80285	Modified factor VI
861	6	3.4	1212	21	AAV46262	Arabidopsis thalia	934	6	3.4	1598	21	AAV43628	Arabidopsis thalia
862	6	3.4	1213	22	AAU33862	Staphylococcus aur	935	6	3.4	1610	22	AAV30568	A full length huma
863	6	3.4	1215	21	AAV30842	Arabidopsis thalia	936	6	3.4	1612	21	AAV43627	Arabidopsis thalia
864	6	3.4	1217	22	AAU36930	Staphylococcus aur	937	6	3.4	1618	22	AAV85504	Human protein kina
865	6	3.4	1220	22	AAV86160	D. melanogaster GA	938	6	3.4	1632	22	ABE62676	Drosophila melanog
866	6	3.4	1221	22	ABE62615	Drosophila melanog	939	6	3.4	1647	22	ABG10750	Novel human diagno
867	6	3.4	1221	22	AAU38942	Drosophila G-prote	940	6	3.4	1661	18	AAV18670	Factor VIII-dB695-
868	6	3.4	1231	22	ABE59738	Drosophila melanog	941	6	3.4	1665	23	AAE24151	Human kinase (PKI
869	6	3.4	1232	21	AAV26239	Human N-methyl-D-a	942	6	3.4	1665	23	AAO15372	Human myosin light
870	6	3.4	1232	23	AAV47961	Human NMDA recepto	943	6	3.4	1922	22	ABE63631	Drosophila melanog
871	6	3.4	1243	22	ABG21221	Novel human diagno	944	6	3.4	2066	22	AAE10146	Streptomyces nous
872	6	3.4	1268	22	ABE67095	Drosophila melanog	945	6	3.4	2090	22	ABE58261	Drosophila melanog
873	6	3.4	1285	22	ABG16337	Novel human diagno	946	6	3.4	2098	17	AAV86863	Factor-VIII. Homo
874	6	3.4	1291	22	ABE71544	Drosophila melanog	947	6	3.4	2115	21	AAV57849	Plg Factor VIII pr
875	6	3.4	1305	22	ABE59497	Drosophila G-prote	948	6	3.4	2133	19	AAV44133	Sus scrofa factor
876	6	3.4	1305	22	AAU38927	Drosophila G-prote	949	6	3.4	2133	20	AAV31597	Porcine factor VII
877	6	3.4	1305	22	AAV86161	D. melanogaster GA	950	6	3.4	2133	22	AAE11203	Porcine factor VII
878	6	3.4	1318	22	AAV28169	Novel human secret	951	6	3.4	2133	22	AAE11203	Porcine factor VII
879	6	3.4	1323	20	AAV49133	GABA-BR1a*Gqo5 fus	952	6	3.4	2280	22	ABE61650	Drosophila melanog
880	6	3.4	1323	23	AAO15099	Human GABA-BR1a*Gq	953	6	3.4	2304	21	AAV57848	Mouse factor VIII
881	6	3.4	1334	12	AAV14481	LkTA: lacZ fusion p	954	6	3.4	2319	19	AAV53485	Murine factor VIII
882	6	3.4	1334	15	AAV50290	LkTA: lacZ fusion p	955	6	3.4	2319	19	AAV44135	Mus musculus facto
883	6	3.4	1352	21	AAV46261	Arabidopsis thalia	956	6	3.4	2319	20	AAV31596	Mouse factor VIII
884	6	3.4	1355	21	AAV30841	Arabidopsis thalia	957	6	3.4	2319	22	AAE11202	Murine factor VIII
885	6	3.4	1373	21	AAV2190	Human ORFX ORF1954	958	6	3.4	2319	22	AAV50467	Mouse factor VIII
886	6	3.4	1375	23	AAU98028	S. mutans glucosyl	959	6	3.4	2332	8	AAV71726	Factor VIII:C varia

960	6	3.4	2332	8	AAP171727
961	6	3.4	2332	8	AAP171728
962	6	3.4	2332	8	AAP171729
963	6	3.4	2332	14	AAW343257
964	6	3.4	2332	18	AAW333222
965	6	3.4	2332	18	AAW333223
966	6	3.4	2332	18	AAW333224
967	6	3.4	2332	18	AAW333225
968	6	3.4	2332	18	AAW333226
969	6	3.4	2332	19	AAW534283
970	6	3.4	2332	19	AAW44113
971	6	3.4	2332	20	AAW311594
972	6	3.4	2332	21	AAW57847
973	6	3.4	2332	22	AAE10826
974	6	3.4	2332	22	AAE11200
975	6	3.4	2332	22	AAW171902
976	6	3.4	2332	22	AAW50465
977	6	3.4	2332	23	AAU79869
978	6	3.4	2332	23	AAU79870
979	6	3.4	2332	23	AAU79871
980	6	3.4	2332	23	AAU79872
981	6	3.4	2332	23	AAU79873
982	6	3.4	2332	23	AAU79874
983	6	3.4	2332	23	AAU79875
984	6	3.4	2342	18	AAW11422
985	6	3.4	2342	18	AAW11349
986	6	3.4	2342	18	AAW11393
987	6	3.4	2343	18	AAW11335
988	6	3.4	2343	20	AAW80989
989	6	3.4	2343	21	AAW57846
990	6	3.4	2344	18	AAW11432
991	6	3.4	2344	18	AAW11410
992	6	3.4	2344	18	AAW11384
993	6	3.4	2344	18	AAW11361
994	6	3.4	2344	18	AAW11367
995	6	3.4	2344	18	AAW11370
996	6	3.4	2344	18	AAW11331
997	6	3.4	2345	18	AAW11453
998	6	3.4	2345	18	AAW11466
999	6	3.4	2345	18	AAW11441
1000	6	3.4	2345	18	AAW11446

ALIGNMENTS

RESULT 1	
AAR12456	
ID	AAR12456 standard; Protein; 77 AA.
XX	
XX	AAR12456;
XX	
XX	26-JUL-1991 (first entry)
XX	
DE	H. influenzae strain Minna (OMP subtype 1H)-cro-lacZ fusion
DE	protein.
XX	
KW	bacterial meningitis; vaccine; pl gene; T-cell antigen; PRSM793;
KW	outer membrane protein.
XX	
XX	Haemophilus influenzae.
XX	
XX	WO9106652-A.
PN	
XX	
PD	16-MAY-1991.
XX	
XX	31-OCT-1990; 90WO-CA00374.
XX	
XX	31-OCT-1989; 89GB-0024473.
PR	
XX	
PA	(CONN-) CONNAUGHT LAB LTD.
PA	(UNIW) WASHINGTON UNIV ST LOUIS.
XX	

PI Munson RS, Grass S, Chong P, Yang Y, Fahim R, McVerryzP;
PI Klein M;
XX
XX WPI; 1991-164201/22.
XX
XX
PT Outer membrane protein of Haemophilus influenzae type B - used as
PT vaccine against infections, esp. in infants and for diagnosis
XX
XX Disclosure; Fig 5; 33pp; English.
XX
XX Plasmid PRSM793 contains only the 3' portion of the pI gene. The
CC plasmid is derived from PRSM188 which contains the full-length pI
CC sequence. This cro-lacZ-ompP1 fusion protein is produced by PRSM793
CC and is recognised by rabbit and guinea pig pI-specific antisera in
CC immunoblot analyses.
CC See also AAR12446-R12455.
XX
XX SQ Sequence 77 AA;

Query Match 7.3%; Score 13; DB 12; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 15 GATYKFTPNLSVD 27

RESULT 2
AAR40060
ID AAR40060 standard; peptide; 35 AA.
XX
AC AAR40060;
XX
XX 04-FEB-1994 (first entry)
XX
DE H1b OMP pI peptide H1bP1-8 (279-312).
XX
XX Haemophilus influenzae; type b; Hib; outer membrane protein; pI; P2;
KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
KW Immunogen.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 35 /note= "May be absent"
FT
XX
PN W09J15205-A.
XX
XX 05-AUG-1993.
XX
XX 03-FEB-1993; 93W0-CA00041.
XX
XX 03-FEB-1992; 92GB-0002219.
XX
XX (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Kandil A, Klein MH, Sia C;
XX
XX WPI; 1993-258681/32.

Synthetic Haemophilus influenzae conjugate vaccine - comprising T-helper cell determinants and B-cell epitope(s) linked to synthetic oligo:saccharide(s)

Table 1; Page 47; 99pp; English.

The sequences given in AAR40053-101 are peptide fragments derived from the Haemophilus influenzae type b (Hib) outer membrane proteins P1, P2 and P6. These peptides may be used in a vaccine against Hib infection and antibodies against these peptides may be used in test kits to detect H. influenzae in a sample. The vaccine may further

CC comprise a immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.

XX Sequence 35 AA;

Query Match 5.6%; Score 10; DB 14; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
 |||||
 Db 2 ELSGFHQLTD 11

RESULT 3

AAR12455
 ID AAR12455 standard; Protein; 36 AA.

XX AC AAR12455;

XX 26-JUL-1991 (first entry)

XX H1BP1-14(400-437) 6U H. influenzae Type b OMP epitope.

XX bacterial meningitis; vaccine; PI gene; T-cell antigen;
 KW outer membrane protein.

XX Haemophilus influenzae.

XX WO9106652-A.

XX 16-MAY-1991.

XX 31-OCT-1990; 90WO-CA00374.

XX 31-OCT-1989; 89GB-0024473.

XX (CONN-) CONNAUGHT LAB LTD.

XX (UNIW) WASHINGTON UNIV ST LOUIS.

XX Munson RS, Grass S, Chong P, Yang Y, Fahim R, McVerry P;
 PI Klein M;

XX WPI; 1991-164201/22.

XX Outer membrane protein of Haemophilus influenzae type B - used as
 PT vaccine against infections, esp. in infants and for diagnosis

XX Disclosure; Table 2; 33pp; English.

XX Synthetic peptide corresponding to amino acids 400 to 437 of the H.
 CC influenzae strain 8358 (OMP subtype 6U) PI protein. The N-terminal
 CC Cys residue is additional and allows coupling to a carrier in one
 CC specific orientation. The free peptide was used as an immunogen to
 CC produce antibodies showing that the peptide comprises both T-helper
 CC determinant and B-cell epitope(s).
 CC See also AAR12446-R12454 and AAR12083.

XX Sequence 36 AA;

Query Match 5.6%; Score 10; DB 12; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
 |||||
 Db 25 ANLYGLNLNY 34

RESULT 4

AAR40066
 ID AAR40066 standard; peptide; 36 AA.

XX AAR40066;

XX 04-FEB-1994 (first entry)

XX Hib OMP PI peptide H1BP1-14 (400-437).

XX Haemophilus influenzae; type b; Hib; outer membrane protein; PI; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "May be absent"

FT Misc-difference 5..14 /note= "Residues are different to those found in PI
 protein of H. influenzae strain IH"

XX WO9315205-A.

XX 05-AUG-1993.

XX 03-FEB-1993; 93WO-CA00041.

XX 03-FEB-1992; 92GB-0002219.

XX (CONN-) CONNAUGHT LAB LTD.

XX Chong P, Kandil A, Klein MH, Sia C;

XX WPI; 1993-258681/32.

XX Synthetic Haemophilus influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)

XX Table 1; Page 48; 99pp; English.

XX The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus influenzae type b (Hib) outer membrane proteins PI,
 CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise a immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.

XX Sequence 36 AA;

Query Match 5.6%; Score 10; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGLNLNY 177
 |||||
 Db 25 ANLYGLNLNY 34

RESULT 5

AAR12454
 ID AAR12454 standard; Protein; 39 AA.

XX AC AAR12454;

XX 26-JUL-1991 (first entry)

XX H1BP1-13(400-437) IH H. influenzae Type b OMP epitope.

XX bacterial meningitis; vaccine; PI gene; T-cell antigen;
 KW outer membrane protein.

OS Haemophilus Influenzae.
 PN WO9106652-A.
 PD 16-MAY-1991.
 XX 31-OCT-1990; 90WO-CA00374.
 XX 31-OCT-1989; 89GB-0024473.
 XX (CONN-) CONNAUGHT LAB LTD.
 PA (UNITW) WASHINGTON UNIV ST LOUIS.
 XX Munson RS, Grass S, Chong P, Yang Y, Fahim R, McVerry P;
 PI Klein M;
 XX WPI; 1991-164201/22.
 DR Outer membrane protein of Haemophilus Influenzae type B - used as
 PT vaccine against infections, esp. in infants and for diagnosis
 XX
 PS Disclosure; Table 2; 33pp; English.
 XX Synthetic peptide corresponding to amino acids 400 to 437 of the H.
 CC influenzae P1 protein. The N-terminal Cys residue is additional and
 CC allows coupling to a carrier in one specific orientation. The
 CC free peptide was used as an immunogen to produce antibodies showing
 CC that the peptide comprises both T-helper determinant and major
 CC immunodominant B-cell epitope.
 CC See also AAR12446-R12453 and AAR12455 and AAQ12083.
 XX
 XX Sequence 39 AA;
 SQ
 Query Match 5.6%; Score 10; DB 12; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 ANLYGLNLNY 177
 DB |||||
 DB 28 ANLYGLNLNY 37
 RESULT 6
 AAR40065
 ID AAR40065 standard; peptide; 39 AA.
 XX
 AC AAR40065;
 XX
 DT 04-FEB-1994 (first entry)
 XX
 DE Hib OMP P1 peptide HIBP1-13 (400-437).
 XX
 KW Haemophilus Influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Region 1..15
 FT /label= C-P1
 FT Region 16..53
 FT /label=,CHIBP2
 XX
 PN WO9315205-A.
 XX
 PD 05-AUG-1993.
 XX
 PP 03-FEB-1993; 93WO-CA00041.
 XX
 PR 03-FEB-1992; 92GB-0002219.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Chong P, Kandil A, Klein MH, Sia C;
 XX
 DR WPI; 1993-258681/32.
 XX
 XX Synthetic Haemophilus Influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 XX
 PS Table 11; Page 59; 99pp; English.
 XX
 CC The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus Influenzae type b (Hib) outer membrane proteins P1,
 CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. Influenzae in a sample. The vaccine may further
 CC comprise an immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.
 XX
 XX Sequence 39 AA;
 SQ
 Query Match 5.6%; Score 10; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 ANLYGLNLNY 177
 DB |||||
 DB 28 ANLYGLNLNY 37
 RESULT 7
 AAR40100
 ID AAR40100 standard; peptide; 53 AA.
 XX
 AC AAR40100;
 XX
 DT 04-FEB-1994 (first entry)
 XX
 DE Hib OMP P1-P2 hybrid peptide 2P6-1P13.
 XX
 KW Haemophilus Influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Region 1..15
 FT /label= C-P1
 FT Region 16..53
 FT /label=,CHIBP2
 XX
 PN WO9315205-A.
 XX
 PD 05-AUG-1993.
 XX
 PP 03-FEB-1993; 93WO-CA00041.
 XX
 PR 03-FEB-1992; 92GB-0002219.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Chong P, Kandil A, Klein MH, Sia C;
 XX
 DR WPI; 1993-258681/32.
 XX
 XX Synthetic Haemophilus Influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 XX
 PS Table 11; Page 59; 99pp; English.
 XX
 CC The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus Influenzae type b (Hib) outer membrane proteins P1,

XX WPI; 1993-258681/32.
 XX Synthetic Haemophilus Influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 XX
 PS Table 1; Page 48; 99pp; English.
 XX
 CC The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus Influenzae type b (Hib) outer membrane proteins P1,
 CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. Influenzae in a sample. The vaccine may further
 CC comprise an immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.
 XX
 XX Sequence 39 AA;
 SQ
 Query Match 5.6%; Score 10; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 ANLYGLNLNY 177
 DB |||||
 DB 28 ANLYGLNLNY 37
 RESULT 7
 AAR40100
 ID AAR40100 standard; peptide; 53 AA.
 XX
 AC AAR40100;
 XX
 DT 04-FEB-1994 (first entry)
 XX
 DE Hib OMP P1-P2 hybrid peptide 2P6-1P13.
 XX
 KW Haemophilus Influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Region 1..15
 FT /label= C-P1
 FT Region 16..53
 FT /label=,CHIBP2
 XX
 PN WO9315205-A.
 XX
 PD 05-AUG-1993.
 XX
 PP 03-FEB-1993; 93WO-CA00041.
 XX
 PR 03-FEB-1992; 92GB-0002219.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Chong P, Kandil A, Klein MH, Sia C;
 XX
 DR WPI; 1993-258681/32.
 XX
 XX Synthetic Haemophilus Influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 XX
 PS Table 11; Page 59; 99pp; English.
 XX
 CC The sequences given in AAR40053-101 are peptide fragments derived from
 CC the Haemophilus Influenzae type b (Hib) outer membrane proteins P1,

CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise an immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.
 XX
 XX Sequence 53 AA;

Query Match 5.6%; Score 10; DB 14; Length 53;
 Best Local Similarity 100.0%; Pred. No. 0.024; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

QY 168 ANLYGLNLNY 177
 DB 42 ANLYGLNLNY 51
 |||||

RESULT 8
 AAR40101
 ID AAR40101 standard; peptide; 54 AA.
 XX
 AC AAR40101;
 XX
 DT 04-FEB-1994 (first entry)
 XX
 DE Hib OMP P1-P2 hybrid peptide CP2-1P13.
 XX
 KW Haemophilus influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..16
 FT /label= C-P1
 FT Region 17..53
 FT /label= CHIBP2
 FT
 XX WO9315205-A.
 XX
 XX 05-AUG-1993.
 XX
 XX 03-FEB-1993; 93WO-CA00041.
 XX
 XX 03-FEB-1992; 92GB-0002219.
 XX
 XX (CONN-) CONNAUGHT LAB LTD.
 XX
 XX Chong P, Kandil A, Klein MH, Sia C;
 XX WPI; 1993-258681/32.
 XX
 XX Synthetic Haemophilus influenzae conjugate vaccine - comprising
 XX T-helper cell determinants and B-cell epitope(s) linked to
 XX synthetic oligo:saccharide(s)
 XX
 XX Table 11; Page 59; 99pp; English.
 XX
 XX The sequences given in AAR40053-101 are peptide fragments derived from
 XX the Haemophilus influenzae type b (Hib) outer membrane proteins P1,
 XX P2 and P6. These peptides may be used in a vaccine against Hib
 XX infection and antibodies against these peptides may be used in test
 XX kits to detect H. influenzae in a sample. The vaccine may further
 XX comprise an immunogenic or immunostimulatory molecule or the peptides
 XX may be modified with lipids, or linked to synthetic PRP as synthetic
 XX lipoglycopeptide conjugates to produce alternative vaccines.
 XX
 XX Sequence 54 AA;

Query Match 5.6%; Score 10; DB 14; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 ANLYGLNLNY 177
 DB 43 ANLYGLNLNY 52
 |||||

RESULT 9
 AAR40099
 ID AAR40099 standard; peptide; 55 AA.
 XX
 AC AAR40099;
 XX
 DT 04-FEB-1994 (first entry)
 XX
 DE Hib OMP P1-P2 hybrid peptide 2P2-1P13.
 XX
 KW Haemophilus influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..17
 FT /label= C-P1
 FT Region 18..55
 FT /label= CHIBP2
 FT
 XX WO9315205-A.
 XX
 XX 05-AUG-1993.
 XX
 XX 03-FEB-1993; 93WO-CA00041.
 XX
 XX 03-FEB-1992; 92GB-0002219.
 XX
 XX (CONN-) CONNAUGHT LAB LTD.
 XX
 XX Chong P, Kandil A, Klein MH, Sia C;
 XX WPI; 1993-258681/32.
 XX
 XX Synthetic Haemophilus influenzae conjugate vaccine - comprising
 XX T-helper cell determinants and B-cell epitope(s) linked to
 XX synthetic oligo:saccharide(s)
 XX
 XX Table 11; Page 59; 99pp; English.
 XX
 XX The sequences given in AAR40053-101 are peptide fragments derived from
 XX the Haemophilus influenzae type b (Hib) outer membrane proteins P1,
 XX P2 and P6. These peptides may be used in a vaccine against Hib
 XX infection and antibodies against these peptides may be used in test
 XX kits to detect H. influenzae in a sample. The vaccine may further
 XX comprise an immunogenic or immunostimulatory molecule or the peptides
 XX may be modified with lipids, or linked to synthetic PRP as synthetic
 XX lipoglycopeptide conjugates to produce alternative vaccines.
 XX
 XX Sequence 55 AA;

Query Match 5.6%; Score 10; DB 14; Length 55;
 Best Local Similarity 100.0%; Pred. No. 0.025;

QY 168 ANLYGLNLNY 177
 DB 44 ANLYGLNLNY 53
 |||||

RESULT 10
 AAB96665
 ID AAB96665 standard; Protein; 636 AA.
 XX

```

AC AAB96665;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssal ATPase subunit of ABC transporter #18.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-AL.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Priour D, Dietrich J, Lecompte O;
PI Querellou J, Welssenbach J, Saurin W, Heilig R;
XX
DR WPI; 2001-126236/14.
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
PT
PS Claim 7; Pages 1419-1421; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF6431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 636 AA;
Query Match 4.5%; Score 8; DB 22; Length 636;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 KKHFFVET 149
DB 322 KKHFFVET 329
|||||||

RESULT 11
AAY33535
ID AAY33535 standard; Protein; 61 AA.
XX
AC AAY33535;
XX
DT 20-JAN-2000 (first entry)
XX
DE Tobacco DNA binding protein EREBP2 AP2 domain.
XX
KW Environmental stress tolerance; plant; binding protein; DNA regulation;
KW CBF-1; C-repeat/DRE binding factor; CCG regulatory sequence; COR; cold;
KW cold-related gene; drought; high salinity; tissue-specific promoter;
KW AP2 domain; tobacco; EREBP2.
XX
OS Nicotiana tabacum.
XX
PN WO9938977-A2.
XX
PD 05-AUG-1999.

```

```

XX 28-JAN-1999; 99WO-US01895.
XX
XX 03-FEB-1998; 98US-0017575.
PR 03-FEB-1998; 98US-0017816.
PR 03-FEB-1998; 98US-0018227.
PR 03-FEB-1998; 98US-0018233.
PR 03-FEB-1998; 98US-0018234.
PR 03-FEB-1998; 98US-0018235.
PR 23-NOV-1998; 98US-0198119.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Stockinger EJ, Jaglo-Ottosen K, Zarka D, Gilmour SJ, Jiang C;
PI Fromm M, Thomashow MF;
XX
DR WPI; 1999-561312/47.
XX
XX Environmental stress tolerance gene binding proteins useful for
PT altering plant stress tolerance -
PT
XX Example 1D; Fig 2D; 252pp; English.
XX
CC This invention describes novel binding proteins other than CBF-1
CC (C-repeat/DRE binding factor) in isolated form which comprise a
CC consensus sequence capable of binding to a CCG regulatory sequence.
CC The binding proteins are capable of binding to a DNA regulatory
CC sequence, which regulates expression of one or more environmental
CC stress tolerance genes, especially COR (cold-related) genes.
CC Environmental stress may be, e.g. cold temperatures, drought and high
CC salinity. Plants transformed with the binding protein (or sequences
CC encoding it) can have altered environmental stress tolerance. The
CC binding protein coding sequences can be under the control of
CC tissue-specific promoters. This sequence represents the tobacco DNA
CC binding protein EREBP2 AP2 domain which is used to describe the method
CC of the invention.
XX
SQ Sequence 61 AA;
Query Match 3.9%; Score 7; DB 20; Length 61;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
DB 40 LAYDKAA 46
|||||

RESULT 12
AAY31478
ID AAY31478 standard; peptide; 68 AA.
XX
AC AAY31478;
XX
DT 29-OCT-1999 (first entry)
XX
DE A. thaliana RAP2.4 EREBP-like subclass AP2 domain.
XX
KW ADC gene; AP2 domain containing gene; regulatory gene; APETALA2;
KW seed mass modulation; genetic engineering; transgenic plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9941974-A1.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US03429.
PR 19-FEB-1998; 98US-0026039.
XX
XX (REGC ) UNIV CALIFORNIA.

```

XX Jofuku KD, Okamuro JK;
 XX WPI; 1999-518486/43.
 XX Novel methods for modulating seed mass and other plant traits using
 XX new expression cassettes containing a plant promoter
 XX Example 4; Page 84; 104pp; English.
 XX The invention provides novel methods for controlling seed size and total
 CC seed protein using ADC (AP2 domain containing) gene, which is a plant
 CC regulatory gene, over expression and antisense gene constructs. The
 CC method of modulating seed mass in a plant comprises: providing a first
 CC plant comprising a recombinant expression cassette containing an ADC
 CC nucleic acid linked to a plant promoter; selfing the first plant or
 CC crossing the first plant with a second plant, thereby producing a
 CC plurality of seeds; and selecting seeds with altered mass. The methods
 CC of the invention can be used to enhance or increase endogenous gene
 CC expression. Enhanced ADC expression leads to smaller seeds or seedless
 CC fruit. The methods can be used to produce a broad range of transgenic
 CC plants. Increasing seed size, amino acid content, and oil content is
 CC desirable in crop plants for human or animal consumption, e.g. soybean,
 CC rice, wheat, corn, rye, etc.. Decreasing seed size is useful in plants
 CC grown for their fruit and where large seeds are undesirable, e.g.
 CC cucumbers, tomatoes, cherries, melons. Sequences AAY31456-496 represent
 CC AP2 domain fragments of various polypeptides.
 XX
 SQ Sequence 68 AA;
 Query Match 3.9%; Score 7; DB 20; Length 68;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 LAYDKAA 101
 Db 41 LAYDKAA 47
 RESULT 13
 AAB25837
 ID AAB25837 standard; Protein; 68 AA.
 XX
 AC AAB25837;
 XX
 DT 15-DEC-2000 (first entry)
 XX
 DE AP2 domain amino acid sequence of Arabidopsis RAP2 protein.
 XX
 KW Apetela; AP2 domain; canola; soybean; ANT; RAP; EREBP; flower; seedless;
 KW fruit; cherry; melon; tomato; AP2 domain containing; ADC;
 KW transgenic plant.
 XX
 OS Arabidopsis sp.
 XX
 PN US6093874-A.
 XX
 PD 25-JUL-2000.
 XX
 PF 15-AUG-1997; 97US-0912272.
 XX
 PR 20-AUG-1996; 96US-0700152.
 PR 20-JUN-1997; 97US-0879827.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Okamuro JK, Jofuku KD;
 XX
 DR WPI; 2000-514122/46.
 XX
 XX Modulating seed traits useful for creating transgenic plants with
 PT altered seed size or protein content, by providing a plant with a
 PT recombinant expression cassette containing an ADC (AP2 domain

PT containing) nucleic acid linked to a promoter -
 XX Disclosure; Figure 4; 60pp; English.
 XX This invention relates to a method for modulating seed traits in a
 CC soybean or canola plant comprising providing a plant having a
 CC recombinant expression cassette containing an ADC (AP2 domain containing)
 CC nucleic acid linked to a plant promoter. ADC gene sequences are
 CC represented by AAA96979-A96981. The AP2 gene (APETALA) was originally
 CC isolated from Arabidopsis sp. and AP2 domain containing genes include
 CC RAP2 (related to AP2) ANT and EREBPs. The AP2 gene is a floral homeotic
 CC gene. Each of these proteins contain AP2 like domains. Sequences
 CC AAB25813-B25858 represent protein fragments and peptides from ADC
 CC proteins such as RAP2.7. Sequences AAA96982-A97035 represent PCR primers
 CC used to isolate the ADC genes of the invention. The method is used for
 CC modulating seed traits in canola and soybean plants and for creating
 CC transgenic plants. The nucleic acids can be used to confer desired traits
 CC on essentially any plant by increasing or decreasing gene expression.
 CC These may be used to increase seed size, protein, amino acid or oil
 CC content in crop plants in which seeds are used directly for animal or
 CC human consumption or for industrial purposes, such as soybean, canola and
 CC other grains. These may be also be used to decrease seed size or
 CC producing seedless varieties in plants grown for their fruit and in which
 CC large seeds may be undesirable, such as cherries, melons or tomatoes.
 CC Primers, which specifically amplify AP2 domains of the genes, are
 CC particularly useful for identification of particular ADC
 CC polynucleotides.
 XX
 SQ Sequence 68 AA;
 Query Match 3.9%; Score 7; DB 21; Length 68;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 LAYDKAA 101
 Db 41 LAYDKAA 47
 RESULT 14
 AAY31468
 ID AAY31468 standard; peptide; 69 AA.
 XX
 AC AAY31468;
 XX
 DT 29-OCT-1999 (first entry)
 XX
 DE N. tabacum EREBP-2 EREBP-like subclass AP2 domain.
 XX
 KW ADC gene; AP2 domain containing gene; regulatory gene; APETALA2;
 KW seed mass modulation; genetic engineering; transgenic plant.
 XX
 OS Nicotiana tabacum.
 XX
 PN WO9941974-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 17-FEB-1999; 99WO-US03429.
 XX
 PR 19-FEB-1998; 98US-0026039.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Jofuku KD, Okamuro JK;
 XX
 DR WPI; 1999-518486/43.
 XX
 XX Novel methods for modulating seed mass and other plant traits using
 PT new expression cassettes containing a plant promoter
 XX Example 4; Page 79-80; 104pp; English.
 XX

CC The invention provides novel methods for controlling seed size and total
 CC seed protein using ADC (AP2 domain containing) gene, which is a plant
 CC regulatory gene, over expression and antisense gene constructs. The
 CC method of modulating seed mass in a plant comprises: providing a first
 CC plant comprising a recombinant expression cassette containing an ADC
 CC nucleic acid linked to a plant promoter; selfing the first plant or
 CC crossing the first plant with a second plant, thereby producing a
 CC plurality of seeds; and selecting seeds with altered mass. The methods
 CC of the invention can be used to enhance or increase endogenous gene
 CC expression. Enhanced ADC expression leads to smaller seeds or seedless
 CC fruit. The methods can be used to produce a broad range of transgenic
 CC plants. Increasing seed size, amino acid content, and oil content is
 CC desirable in crop plants for human or animal consumption, and oil content is
 CC desirable in rice, wheat, corn, rye, etc.. Decreasing seed size is useful in plants
 CC grown for their fruit and where large seeds are undesirable, e.g.
 CC cucumbers, tomatoes, cherries, melons. Sequences AAB31456-496 represent
 CC AP2 domain fragments of various polypeptides.

XX SQ Sequence 69 AA;

Query Match 3.9%; Score 7; DB 20; Length 69;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
 Db 42 LAYDKAA 48
 |||||

RESULT 15

AAB25827
 ID AAB25827 standard; Protein: 69 AA.

XX AC AAB25827;

XX DT 15-DEC-2000 (first entry)

XX DE AP2 domain amino acid sequence of tobacco EREBP.

XX KW Apetela; AP2 domain; canola; soybean; ANT; RAP; EREBP; flower; seedless;
 KW fruit; cherry; melon; tomato; AP2 domain containing; ADC;
 KW transgenic plant.

XX OS Nicotiana acuminata.

XX PN US6093874-A.

XX PD 25-JUL-2000.

XX PF 15-AUG-1997; 97US-0912272.

XX PR 20-AUG-1996; 96US-0700152.

XX PR 20-JUN-1997; 97US-0879827.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Okamuro JK, Jofuku KD;

XX WPI; 2000-514122/46.

XX Modulating seed traits useful for creating transgenic plants with
 PT altered seed size or protein content, by providing a plant with a
 PT recombinant expression cassette containing an ADC (AP2 domain
 PT containing) nucleic acid linked to a promoter.

XX Disclosure; Figure 4; 60pp; English.

XX This invention relates to a method for modulating seed traits in a
 CC soybean or canola plant comprising providing a plant having a
 CC recombinant expression cassette containing an ADC (AP2 domain containing)
 CC nucleic acid linked to a plant promoter. ADC gene sequences are
 CC represented by AAA96979-A96981. The AP2 gene (APETALA) was originally
 CC isolated from Arabidopsis sp. and AP2 domain containing genes include

CC RAP2 (related to AP2) ANT and EREBPs. The AP2 gene is a floral homeotic
 CC gene. Each of these proteins contain AP2 like domains. Sequences
 CC AAB25813-B25858 represent protein fragments and peptides from ADC
 CC proteins such as RAP2.7. Sequences AAA96982-A97035 represent PCR primers
 CC used to isolate the ADC genes of the invention. The method is used for
 CC modulating seed traits in canola and soybean plants and for creating
 CC transgenic plants. The nucleic acids can be used to confer desired traits
 CC on essentially any plant by increasing or decreasing gene expression.
 CC These may be used to increase seed size, protein, amino acid or oil
 CC content in crop plants in which seeds are used directly for animal or
 CC human consumption or for industrial purposes, such as soybean, canola and
 CC other grains. These may be also be used to decrease seed size or
 CC producing seedless varieties in plants grown for their fruit and in which
 CC large seeds may be undesirable, such as cherries, melons or tomatoes.
 CC Primers, which specifically amplify AP2 domains of the genes, are
 CC particularly useful for identification of particular ADC
 CC polynucleotides.

XX SQ Sequence 69 AA;

Query Match 3.9%; Score 7; DB 21; Length 69;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
 Db 42 LAYDKAA 48
 |||||

RESULT 16

AAG32799
 ID AAG32799 standard; Protein: 79 AA.

XX AC AAG32799;

XX DT 17-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 39638.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 29-MAR-1999; 99US-0126264.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 3.9%; Score 7; DB 21; Length 79;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 151 NIKGLLL 157
Db 5 NIKGLLL 11
RESULT 17
AAG92876
ID AAG92876 standard; Protein: 93 AA.
XX AC AAG92876;
XX XX
DT 26-SEP-2001 (first entry)
XX XX
DE C glutamicum protein fragment SEQ ID NO: 6630.
XX XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX XX
OS Corynebacterium glutamicum.
XX XX
PN EPI108790-A2.
XX XX
PD 20-JUN-2001.
XX XX
PF 18-DEC-2000; 2000EP-0127688.
XX XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochial K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX XX
WPI: 2001-376931/40.
N-PSDB: AAH68095.
XX XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX XX
PS Claim 17; SEQ ID NO: 6630; 246pp + Sequence Listing; English.
XX XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX XX
SQ Sequence 93 AA;
Query Match 3.9%; Score 7; DB 22; Length 93;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 IKGLLLV 158
Db 82 IKGLLLV 88
RESULT 18
AAG32798
ID AAG32798 standard; Protein: 95 AA.
XX AC AAG32798;
XX XX
DT 17-OCT-2000 (first entry)
XX XX
DE Zea mays protein fragment SEQ ID NO: 39637.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX XX
OS Zea mays subsp. mays.
XX XX
PN EPI033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159294.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		

Query Match 3.9%; Score 7; DB 21; Length 95;
Best Local Similarity 100.0%; Pred. No. 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 NIKGLLL 157
|||||
Db 21 NIKGLLL 27

RESULT 19
AAU54058
ID AAU54058 standard; Protein: 109 AA.
XX
AC AAU54058;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes Immunogenic protein #14954.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN W0200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
WPI: 2001-616774/71.
DR N-PSDB; AAS59563.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID No 15253; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 109 AA;

Query Match 3.9%; Score 7; DB 22; Length 109;
Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 LSASIPD 112
|||||
Db 68 LSASIPD 74

RESULT 20
AAU48379
ID AAU48379 standard; Protein: 113 AA.
XX
AC AAU48379;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes Immunogenic protein #9275.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN W0200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
WPI: 2001-616774/71.
DR N-PSDB; AAS59542.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID No 9574; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 113 AA;

Query Match 3.9%; Score 7; DB 22; Length 113;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 RAGLAYD 98
|||||||
Db 45 RAGLAYD 51

RESULT 21
ABP25758
ID ABP25758 standard; Protein; 127 AA.
XX
AC ABP25758;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 692.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maslignani V, Margarit Ros VI, Grandi G, Fraser C;
PI Tetteilin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN66389.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein.
XX
PS Claim 1; Page 3221; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
XX Sequence 127 AA;

Query Match 3.9%; Score 7; DB 23; Length 127;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 SKTHLSA 108

Db 101 SKTHLSA 107
|||||||
RESULT 22
AAG47231
ID AAG47231 standard; Protein; 134 AA.
XX
AC AAG47231;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59507.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 05-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.9%; Score 7; DB 21; Length 134;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 IKGLLV 158
|||||
Db 2 IKGLLV 8

RESULT 23
AAW60957
ID AAW60957 standard; Protein: 206 AA.

XX AAW60957;
 XX 13-OCT-1998 (first entry)
 XX Streptococcus pneumoniae encoded polypeptide.
 XX coding region; ORF; open reading frame; antibacterial;
 KW infection; prevention; meningitis.
 XX Streptococcus pneumoniae.
 XX WO9819689-A1.
 XX 14-MAY-1998.
 XX 27-OCT-1997; 97WO-US19226.
 XX 01-NOV-1996; 96US-0029930.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Reid RH, Zarfos PN;
 XX WPI; 1998-286586/25.
 DR N-PSDB; AAV37356.
 XX New isolated nucleic acids from Streptococcus pneumoniae - useful,
 PT e.g. for identifying anti-bacterial(s) for treatment and prevention
 PT of meningitis
 XX Claim 11; Page 57; 130pp; English.
 CC The sequence is that of the polypeptide encoded by a region isolated
 CC from S. pneumoniae which shows homology to ISL2 protein.
 CC The protein, or agonists of it, may be useful as an antibacterial for
 CC treatment or prevention of infection, specifically caused by S.
 CC pneumoniae (particularly meningitis) but possibly also Helicobacter
 CC pylori (ulcers and gastric cancer). It may be of particular
 CC use before insertion of an in-dwelling device or any other
 CC invasive procedure. The protein, or nucleic acid encoding
 CC it, can also be used in vaccines to induce a cellular
 CC and/or humoral immune response, or to screen for other
 CC antibacterials. The DNA may also contain flanking sequences
 CC that are potential sources of control elements for bacterial
 CC gene expression. Detecting a sequence encoding the protein
 CC can be used diagnostically, e.g. to detect a mutation for
 CC serotyping or classifying infectious agents.
 XX Sequence 206 AA;
 SQ Query Match 3.9%; Score 7; DB 19; Length 206;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 102 SKTHLSA 108
 Db | | | | | | | |
 40 SKTHLSA 46
 RESULT 24
 ABP26302
 ID: ABP26302 standard; Protein; 207 AA.
 XX AC ABP26302;
 XX 02-JUL-2002 (first entry)
 XX Streptococcus polypeptide SEQ ID NO 1780.
 DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW

KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.
 XX WO200234771-A2.
 XX 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-GB04789.
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN66933.
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein.
 XX Claim 1; Page 3328; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX Sequence 207 AA;
 SQ Query Match 3.9%; Score 7; DB 23; Length 207;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 102 SKTHLSA 108
 Db | | | | | | | |
 181 SKTHLSA 187
 RESULT 25
 AAE02538
 ID: AAE02538 standard; Protein; 218 AA.
 XX AC AAE02538;
 XX 10-AUG-2001 (first entry)
 XX A. thaliana transcription factor G26.
 DE Plant transcription factor; phenotype; sugar sensing characteristic;
 KW transgenic plant; plant yield; growth; germination; photosynthesis;
 KW glyoxylate metabolism; respiration; pathogen response; wounding response;
 KW cell cycle regulation; pigmentation; flowering; senescence; physiology;

KW storage organ; metabolism.
 XX Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT 67..134
 FT /note= "Conserved domain"
 XX
 XX WO200135725-A1.
 XX
 XX 25-MAY-2001.
 XX
 XX 14-NOV-2000; 2000WO-US31414.
 XX
 XX 17-NOV-1999; 99US-0166228.
 XX 17-APR-2000; 2000US-0197899.
 XX 22-AUG-2000; 2000US-0227439.
 XX
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J.
 PA (PINE/) PINEDA O.
 PA (PILG/) PILGRIM M.
 PA (ADAM/) ADAM L.
 PA (RIEC/) RIECHMANN J L.
 PA (YUGG/) YU G.
 PA (SAMA/) SAMAHA R.
 XX
 XX JIANG C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;
 PI Yu G, Samaha R;
 XX
 XX WPI; 2001-335977/35.
 XX N-PSDB; AAD06639.
 XX
 XX Nucleic acids encoding plant transcription factor polypeptides, useful
 PT for altering the sugar sensing characteristics of plants and increasing
 PT yield, e.g. corn, potato and cotton plants -
 XX
 XX Claim 4; Page 57-58; 151pp; English.
 XX
 CC The patent relates to polynucleotides encoding 35 plant transcription
 CC factors which may be used to modify phenotype associated with a plant's
 CC sugar sensing characteristics and increasing yield when their expression
 CC level is altered. Sugars are central regulatory molecules that control
 CC aspects of physiology, metabolism and development. Therefore the cDNAs
 CC and proteins of the invention are useful for modifying the growth and
 CC germination rates of plants, photosynthesis, glyoxylate metabolism,
 CC respiration, starch and sucrose synthesis and degradation, pathogen
 CC response, wounding response, cell cycle regulation, pigmentation,
 CC flowering and senescence of plants and for modifying sink-source
 CC relationships in seeds, tubers, roots, and other storage organs leading
 CC to an increase in yield. The transcription factor polynucleotides and
 CC polypeptides may be used to alter the structure and developmental
 CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
 CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
 CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
 CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
 CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
 CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.
 CC The present sequence is an Arabidopsis thaliana transcription factor.
 XX
 SQ Sequence 218 AA;
 Query Match 3.9%; Score 7; DB 22; Length 218;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 LAYDKAA 101
 |||||
 Db 107 LAYDKAA 113
 RESULT 26

AAB32762
 ID AAB32762 standard; Protein; 219 AA.
 XX
 AC AAB32762;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE Eucalyptus grandis transcription factor protein sequence #220.
 XX
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.
 XX
 OS Eucalyptus grandis.
 XX
 PN WO200053724-A2.
 XX
 XX 14-SEP-2000.
 XX
 XX 09-MAR-2000; 2000WO-US06112.
 XX
 XX 11-MAR-1999; 99US-0266513.
 XX 18-AUG-1999; 99US-0149485.
 PR
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 XX Wood M, McGrath A, Shenk MA, Glenn M;
 PI
 XX WPI; 2000-579369/54.
 XX
 XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT -
 XX
 XX Claim 8; Page 293; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.
 XX
 SQ Sequence 219 AA;
 Query Match 3.9%; Score 7; DB 21; Length 219;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 LAYDKAA 101
 |||||
 Db 65 LAYDKAA 71
 RESULT 27
 AAY21977
 ID AAY21977 standard; Protein; 230 AA.
 XX
 AC AAY21977;
 XX
 XX 06-SEP-1999 (first entry)
 DT
 XX Sark protein fragment.
 DE
 XX

KW Senescence-associated receptor-like protein kinase; sark; sam gene; sag;
 KW senescence-associated gene; plant senescence; promoter; pharmaceutical;
 KW plant maturation; S-adenosyl methionine; flower; fruit development.

XX Phaseolus vulgaris.

XX WO9929159-A1.

XX 17-JUN-1999.

XX 08-DEC-1998; 98WO-US25799.

XX 08-DEC-1997; 97US-0067898.

XX (VITA-) VITALITY BIOTECHNOLOGIES INC.

XX Cepstein S, Hajjuje T, Rosner A;

XX WPI; 1999-404873/34.

XX N-PSDB; AAX81126.

XX DNA encoding senescence-associated genes for a senescence
 PT receptor-like protein kinase

PS Disclosure; Fig 2; 70pp; English.

XX The invention relates to a senescence-associated receptor-like protein
 CC kinase (sark) gene. The sark gene is a senescence-associated gene (sag)
 CC and is expressed early in the plant senescence process. The sark gene
 CC promoter is useful for driving expression of foreign genes having a
 CC desired product, such as a pharmaceutical, during the process of plant
 CC maturation. The sark gene promoter can be used to drive expression of
 CC resistance genes against pathogens or pests during senescence when the
 CC plant is particularly susceptible to infection or infestation. The sark
 CC gene promoter may also be used to drive expression of a gene encoding an
 CC inhibitor of senescence. Plant senescence may be inhibited by use of
 CC antisense sark constructs. Over expression of the sag genes, using the
 CC sark or sam (S-adenosyl methionine) gene promoters is useful for
 CC induction of early senescence. This is useful to obtain flower or fruit
 CC development prior to specific pest onset, prior to undesirable cross-
 CC fertilization from related crops, at a specific time during storage or
 CC retail, or to avoid development of plant structures that are not of
 CC agronomic importance. The present sequence represents a region of the
 CC sark protein selected for expression in E. coli.

SQ Sequence 230 AA;

Query Match 3.9%; Score 7; DB 20; Length 230;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 LSASIPD 112

Db 144 LSASIPD 150

RESULT 28

AAG38454

ID AAG38454 standard; Protein; 313 AA.

XX AC AAG38454;

XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 47443.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PN

XX 06-SEP-2000. 2000EP-0301439.
 PD 99US-0121825.
 XX 99US-0123180.
 XX 99US-0123548.
 XX 99US-0125788.
 XX 99US-0126264.
 XX 99US-0126785.
 XX 99US-0127462.
 XX 99US-0128234.
 XX 99US-0128714.
 XX 99US-0129845.
 XX 99US-0130077.
 XX 99US-0130449.
 XX 99US-0130510.
 XX 99US-0130891.
 XX 99US-0131449.
 XX 99US-0132048.
 XX 99US-0132407.
 XX 99US-0132484.
 XX 99US-0132485.
 XX 99US-0132486.
 XX 99US-0132487.
 XX 99US-0132863.
 XX 99US-0134256.
 XX 99US-0134218.
 XX 99US-0134219.
 XX 99US-0134221.
 XX 99US-0134370.
 XX 99US-0134768.
 XX 99US-0134941.
 XX 99US-0135124.
 XX 99US-0135353.
 XX 99US-0135629.
 XX 99US-0136021.
 XX 99US-0136392.
 XX 99US-0136782.
 XX 99US-0137222.
 XX 99US-0137528.
 XX 99US-0137502.
 XX 99US-0137724.
 XX 99US-0138094.
 XX 99US-0138540.
 XX 99US-0138847.
 XX 99US-0139119.
 XX 99US-0139452.
 XX 99US-0139453.
 XX 99US-0139452.
 XX 99US-0139454.
 XX 99US-0139455.
 XX 99US-0139456.
 XX 99US-0139457.
 XX 99US-0139458.
 XX 99US-0139459.
 XX 99US-0139460.
 XX 99US-0139461.
 XX 99US-0139462.
 XX 99US-0139463.
 XX 99US-0139750.
 XX 99US-0139763.
 XX 99US-0139817.
 XX 99US-0139899.
 XX 99US-0140353.
 XX 99US-0140354.
 XX 99US-0140695.
 XX 99US-0140823.
 XX 99US-0140991.
 XX 99US-0141287.
 XX 99US-0141842.
 XX 99US-0142154.

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 3.9%; Score 7; DB 21; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 95 LAYDKAA 101
| | | | |
Db 168 LAYDKAA 174

RESULT 29

AAG38787
ID AAG38787 standard; Protein; 313 AA.

XX AC AAG38787;
XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47902.
XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135355.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.9%; Score 7; DB 21; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 LAYDKAA 101
Db 168 LAYDKAA 174
|||||

RESULT 30
AAG38453
ID AAG38453 standard; Protein; 330 AA.
AC AAG38453;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47442.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

101
99US-0129845.
99US-0130077.
99US-0130449.
99US-0130510.
99US-0130891.
99US-0131449.
99US-0132048.
99US-0132407.
99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0134256.
99US-0134218.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
```

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149369.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156599.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.9%; Score 7; DB 21; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
Db 185 LAYDKAA 191
|||||

RESULT 31
AAG38786
ID AAG38786 standard; Protein; 330 AA.
AC AAG38786;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47901.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.


```
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      3.9%; Score 7; DB 21; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 LAYDKAA 101
DB 185 LAYDKAA 191
      |||||
      |||||

RESULT 32
AAG38452
ID AAG38452 standard; Protein; 334 AA.
XX
AC AAG38452;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47441.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
```

PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161922.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 06-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
PR 07-OCT-1999;	99US-0158029.		
PR 08-OCT-1999;	99US-0158232.		
PR 12-OCT-1999;	99US-0158369.		
PR 13-OCT-1999;	99US-0159293.		
PR 13-OCT-1999;	99US-0159294.		
PR 13-OCT-1999;	99US-0159295.		
PR 14-OCT-1999;	99US-0159329.		
PR 14-OCT-1999;	99US-0159330.		
PR 14-OCT-1999;	99US-0159331.		
PR 14-OCT-1999;	99US-0159637.		
PR 14-OCT-1999;	99US-0159638.		
PR 18-OCT-1999;	99US-0159584.		
PR 21-OCT-1999;	99US-0160741.		
PR 21-OCT-1999;	99US-0160767.		
PR 21-OCT-1999;	99US-0160768.		
PR 21-OCT-1999;	99US-0160770.		
PR 21-OCT-1999;	99US-0160814.		
PR 21-OCT-1999;	99US-0160815.		
PR 22-OCT-1999;	99US-0160980.		
PR 22-OCT-1999;	99US-0160981.		

Query Match

Best Local Similarity

Matches

7; Conservative

3.9%; 100.0%;

DB 21; Length 334;

Pred. No. 1.5e+02;

Mismatches 0;

Indels 0;

Gaps 0;

Oy	95 LAYDKAA 101
Db	189 LAYDKAA 195

RESULT 33

AAG38785

ID AAG38785 standard; Protein; 334 AA.

XX

AC AAG38785;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47900.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0156559.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.

PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 3.9%; Score 7; DB 21; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
 |||||
 DB 189 LAYDKAA 195

RESULT 34

AAG71496
 ID AAG71496 standard; Protein; 338 AA.

XX AAG71496;

XX 31-JUL-2001 (first entry)

XX Human olfactory receptor polypeptide, SEQ ID NO: 1177.

XX Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

XX 06-OCT-2000; 2000WO-US27582.

XX 08-OCT-1999; 99US-0158615.

XX 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -

XX Claim 11; Page 693-694; 1857pp; English.

XX The present sequence is an olfactory receptor which is encoded by
 CC one of a number of novel polynucleotides. The polynucleotides can be
 CC used in screening for olfactory agonists and antagonists. The methods
 CC allow for the determination of primary scents and the identification
 CC of the odour receptors used to detect these primary scents. The methods
 CC also enable determination of secondary scents and the identification of
 CC combinations of odour receptors that are involved in detecting such
 CC secondary scents. This enables the construction of a scent representation
 CC (also called a scent fingerprint or scent profile), which may be used to
 CC re-create and edit scents. Libraries of olfactory receptors are useful
 CC for determining the interaction pattern of a composition with the
 CC receptors, and can be used for determining differences in the olfactory
 CC faculties of different individuals.

XX Sequence 338 AA;

Query Match 3.9%; Score 7; DB 22; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SLTLKLP 20
 |||||

DB 167 SLTLKLP 173

RESULT 35

AAG72390
 ID AAG72390 standard; Protein; 338 AA.

XX AAG72390;

XX 31-JUL-2001 (first entry)

XX Human OR-like polypeptide query sequence, SEQ ID NO: 2071.

XX Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

XX 06-OCT-2000; 2000WO-US27582.

XX 08-OCT-1999; 99US-0158615.

XX 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -

XX Example 6; Page 1388-1389; 1857pp; English.

XX The present sequence is a polypeptide encoded by one of 344 newly mined
 CC human genes. It was used as a query sequence in a database search of
 CC olfactory receptor (OR)-like sequences. The invention relates to isolated
 CC polynucleotides encoding polypeptides involved in olfactory sensation.
 CC The polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary scents
 CC and the identification of the odour receptors used to detect these
 CC primary scents. The methods also enable determination of secondary scents
 CC and the identification of combinations of odour receptors that are
 CC involved in detecting such secondary scents. This enables the
 CC construction of a scent representation (also called a scent fingerprint
 CC or scent profile), which may be used to re-create and edit scents.
 CC Libraries of olfactory receptors are useful for determining the
 CC interaction pattern of a composition with the receptors, and can be
 CC used for determining differences in the olfactory faculties of different
 CC individuals.

XX Sequence 338 AA;

Query Match 3.9%; Score 7; DB 22; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SLTLKLP 20
 |||||

DB 167 SLTLKLP 173

RESULT 36

AAW17820
 ID AAW17820 standard; Protein; 466 AA.

XX AAW17820;

XX 29-JUL-1997 (first entry)
 XX Pseudomonas putida amidase.
 DE Amidase; nitrile hydratase; stereospecific enzyme; chiral compound;
 KW amide; carboxylic acid.
 XX Pseudomonas putida strain NRRL 18668.
 OS WO9712964-A2.
 PN 10-APR-1997.
 PD 03-OCT-1996; 96WO-US15969.
 XX 06-OCT-1995; 95US-0004914.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Failon RD, Nelson MJ, Payne MS;
 PI WPI; 1997-226208/20.
 DR N-PSDB; AAT66442.
 XX Nucleic acid encoding stereospecific nitrile hydratase and amidase -
 PT also transformed cells expressing these enzymes, useful for
 PT conversion of racemic nitrile compounds to chiral amide(s) and acids
 XX Claim 33; Page 58-60; 85pp; English.
 XX The amidase (AAW17820) of Pseudomonas putida NRRL 18668 is capable of
 CC converting a racemic mixture of C3 to C6 amides to the corresponding
 CC enantiomeric R- or S-carboxylic acids. Its amino acid sequence was
 CC deduced from a gene sequence (AAT66442) identified upstream of the
 CC 18668 nitrile hydratase gene (AAT66440). The 18668 amidase is
 CC distinct from the amidase isolated from P. putida B23 (FERM B-187).
 CC Co-expression of the 18668 amidase with nitrile hydratase alpha and
 CC beta subunits (see also AAW17816-17) in microbial host cells allows
 CC the conversion of racemic nitrile compounds to the corresponding
 CC enantiomeric R- or S-carboxylic acids.
 XX SQ Sequence 466 AA;
 Query Match 3.9%; Score 7; DB 18; Length 466;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 91 LAGLAY 97
 DB 443 LAGLAY 449
 RESULT 37
 AAW55454
 ID AAW55454 standard; Protein: 511 AA.
 XX AC AAW55454;
 XX 24-JUN-1998 (first entry)
 XX H. pylori ORF 02gell622_875260_f3_36 cell envelope OMP.
 DE Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 XX Helicobacter pylori.
 OS WO9737044-A1.
 PN 09-OCT-1997.
 PD 09-OCT-1997.
 XX

PF 27-MAR-1997; 97WO-US05223.
 XX 06-DEC-1996; 96US-0761318.
 PR 29-MAR-1996; 96US-0625811.
 PR 02-APR-1996; 96US-0758731.
 PR 25-OCT-1996; 96US-0736905.
 PR 28-OCT-1996; 96US-0738859.
 XX (ASTR) ASTRA AB.
 XX Alm RA, Smith D;
 PI WPI; 1997-503122/46.
 DR N-PSDB; AAV24863.
 XX Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 XX Claims 14,80; Page 661-662; 1145pp; English.
 XX This sequence is a H. pylori cell envelope outer membrane protein
 CC (OMP) having a terminal Phe residue.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or expected
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 XX SQ Sequence 511 AA;
 Query Match 3.9%; Score 7; DB 18; Length 511;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 KGSULTK 18
 DB 283 KGSULTK 289
 RESULT 38
 AAW98255
 ID AAW98255 standard; Protein: 511 AA.
 XX AC AAW98255;
 XX 31-MAR-1999 (first entry)
 DE H. pylori GHPO 1559 protein.
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease.
 XX Helicobacter pylori.
 OS WO9843478-A1.
 PN 08-OCT-1998.
 PD 08-OCT-1998.
 XX

PF 01-APR-1998; 98WO-US06371.
 XX 29-JUL-1997; 97US-0902615.
 PR 01-APR-1997; 97US-0833457.
 PR 24-JUN-1997; 97US-0881227.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
 XX WPI; 1998-542293/46.
 DR N-PSDB; AAX13974.
 XX New isolated Helicobacter polynucleotides - used to develop products
 PT for the diagnosis, prevention and treatment of Helicobacter
 PT infections and gastrointestinal diseases
 XX Claim 8; Page 241-244; 2054pp; English.
 XX This sequence represents a Helicobacter pylori GHPO protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastroduodenal diseases associated with
 CC these infections, including acute, chronic, and atrophic gastritis, and
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 XX Sequence 511 AA;

Query Match 3.9%; Score 7; DB 19; Length 511;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 KGSLLTK 18
 |||||
 Db 283 KGSLLTK 289

RESULT 39
 AAW71519
 ID AAW71519 standard; Protein; 511 AA.
 AC AAW71519;
 XX 09-NOV-1998 (first entry)
 DT Helicobacter polypeptide GHPO 1559.
 DE GHPO 1559; Infection; therapy; diagnosis; vaccine; gastritis;
 XX ulcer.
 KW Helicobacter pylori.
 OS WO9821225-A1.
 PN 22-MAY-1998.
 XX 14-NOV-1997; 97WO-US21353.
 XX 29-JUL-1997; 97US-0902615.
 PR 14-NOV-1996; 96US-0749031.
 PR 01-APR-1997; 97US-0831309.
 PR 01-APR-1997; 97US-0833457.
 PR 01-APR-1997; 97US-0834705.
 PR 24-JUN-1997; 97US-0881227.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.
 PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;
 PI Odenbreit S, Tomb J;

XX WPI; 1998-297855/26.
 DR N-PSDB; AAV52054.
 XX Helicobacter polynucleotide and polypeptide sequences - useful to
 PT treat or prevent gastrointestinal infection
 XX Claim 1; Page 208-209; 362pp; English.
 XX This claimed Helicobacter pylori polypeptide, designated GHPO 1559,
 CC can be used in vaccination methods for preventing or treating
 CC Helicobacter infection. 85 Helicobacter polypeptides (see
 CC AAW71474-W71558) are claimed, as well as isolated polynucleotides
 CC (see AAV52009-93) that encode them. The invention also provides:
 CC methods for producing these Helicobacter polypeptides in
 CC recombinant host systems, and related expression cassettes, vectors
 CC and transformed or transfected host cells; live vaccine vectors
 CC that contain the polynucleotides of the invention and which can be
 CC used to prevent or treat Helicobacter infection; therapeutic and/or
 CC prophylactic methods involving administration of polynucleotide
 CC molecules, polypeptides or monospecific antibodies; methods for
 CC detecting the presence of Helicobacter in samples using e.g.
 CC the polypeptides or monospecific antibodies; and methods for
 CC purifying the polypeptides by antibody-based affinity
 CC chromatography.
 XX Sequence 511 AA;

Query Match 3.9%; Score 7; DB 19; Length 511;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 KGSLLTK 18
 |||||
 Db 283 KGSLLTK 289

RESULT 40
 AAY17176
 ID AAY17176 standard; Protein; 511 AA.
 AC AAY17176;
 XX 03-AUG-1999 (first entry)
 DT H. pylori outer membrane polypeptide.
 DE Outer membrane polypeptide; OMP; vaccine; H. pylori infection; humoral;
 XX cellular immune response.
 KW Helicobacter pylori.
 OS WO9921959-A2.
 PN 06-MAY-1999.
 XX 28-OCT-1998; 98WO-US22883.
 XX 17-DEC-1997; 97US-0993001.
 PR 28-OCT-1997; 97US-0959131.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Alm RA, Ellis RW, Guild BC, Noonan BM, Smith D;
 XX WPI; 1999-326698/27.
 DR N-PSDB; AAX75795.
 XX Cellular vaccine against Helicobacter pylori
 PT Claim 7; Page 229-231; 352pp; English.
 XX The invention relates to a vaccine for preventing or treating infections

CC by Helicobacter pylori. The vaccine contains at least one isolated
CC H. pylori polypeptide, or its fragments, in a carrier, where the
CC carrier is a Salmonella, Vibrio cholerae or Shigella vector containing a
CC nucleic acid encoding the H. pylori polypeptide. The vaccines induce
CC humoral and cellular immune responses. The vaccines are used to treat or
CC prevent infections by H. pylori. Sequences AAX75779 to AAX75837 represent
CC nucleic acid sequences encoding H. pylori outer membrane polypeptides
CC (OMPs) AAY17160 to AAY17218.
XX
SQ Sequence 511 AA:

Query Match 3.9%; Score 7; DB 20; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 KGSLLTK 18
 |||||
DB 283 KGSLLTK 289

Search completed: May 12, 2003, 09:47:38
Job time : 61 secs

; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/037,761
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881

; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 184
; LENGTH: 168

Query Match 3.4%; Score 6; DB 9; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LTLKLP 20
| | | | |
Db 43 LTLKLP 48

Search completed: May 12, 2003, 09:56:07
Job time : 41 secs

```
? ZIP: 20850
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
? COMPUTER: HP Vectra 486/33
? OPERATING SYSTEM: MSDOS version 6.2
? SOFTWARE: ASCII Text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/781,986A
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Benson, Bob
? REGISTRATION NUMBER: 30,446
? REFERENCE/DOCKET NUMBER: PB248PP
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (301) 309-8504
? TELEFAX: (301) 309-8512
? INFORMATION FOR SEQ ID NO: 5220:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 162 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-781-986A-5220

Query Match 3.4%; Score 6; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRGKKK 144
Db 96 LRGKKK 101

RESULT 40
US-09-981-876-184
? Sequence 184: Application US/09981876
? Patent No. US20020184669A1
? GENERAL INFORMATION:
? APPLICANT: Rosen et al.
? TITLE OF INVENTION: 70 Human Secreted Proteins
? FILE REFERENCE: P2001P1
? CURRENT APPLICATION NUMBER: US/09/981,876
? CURRENT FILING DATE: 2001-10-19
? PRIOR APPLICATION NUMBER: 09/148,545
? PRIOR FILING DATE: 1998-09-04
? PRIOR APPLICATION NUMBER: 60/040,162
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,333
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/038,621
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,161
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,626
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,334
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,336
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/040,163
? PRIOR FILING DATE: 1997-03-07
? PRIOR APPLICATION NUMBER: 60/047,615
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,600
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,597
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,502
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,633
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,583
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,617
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,618
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,503
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,592
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,581
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,584
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,500
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,587
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,492
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,598
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,613
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,582
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,596
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,612
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/047,632
? PRIOR FILING DATE: 1997-05-23
? PRIOR APPLICATION NUMBER: 60/043,580
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,568
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,314
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,569
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,311
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,671
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,674
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,669
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,312
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,313
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,672
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/043,315
? PRIOR FILING DATE: 1997-04-11
? PRIOR APPLICATION NUMBER: 60/048,974
? PRIOR FILING DATE: 1997-06-06
? PRIOR APPLICATION NUMBER: 60/056,886
? PRIOR FILING DATE: 1997-08-22
? PRIOR APPLICATION NUMBER: 60/056,877
? PRIOR FILING DATE: 1997-08-22
? PRIOR APPLICATION NUMBER: 60/056,889
? PRIOR FILING DATE: 1997-08-22
? PRIOR APPLICATION NUMBER: 60/056,893
? PRIOR FILING DATE: 1997-08-22
? PRIOR APPLICATION NUMBER: 60/056,630
? PRIOR FILING DATE: 1997-08-22
```

```
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US/09/983,802
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,803
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,684
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12

; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (61)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals stop translation
US-09-983-802-196

Query Match 3.4%; Score 6; DB 9; Length 149;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 QDGSY 62
Db 43 QDGSY 48
|||||

RESULT 38
US-09-925-300-1688
; Sequence 1688, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1688
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1688

Query Match 3.4%; Score 6; DB 10; Length 153;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRKKK 144
Db 39 LRKKK 44
|||||

RESULT 39
US-08-781-986A-5220
; Sequence 5220, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
```

; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4318
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4318

Query Match 3.4%; Score 6; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 103 KTHLSA 108
|||||
DB 123 KTHLSA 128

RESULT 34
US-09-862-540-51
; Sequence 51, Application US/09862540
; Publication No. US20030023992A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; TITLE OF INVENTION: No. US20030023992A1 G Protein-Coupled Receptors
; FILE REFERENCE: 00252US1
; CURRENT APPLICATION NUMBER: US/09/862,540
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,138
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/206,139
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/208,976
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-540-51

Query Match 3.4%; Score 6; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LYGLNL 175
|||||
DB 69 LYGLNL 74

RESULT 35
US-10-102-806-799
; Sequence 799, Application US/10102806
; Publication No. US20030054421A1.
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 799
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-799

Query Match 3.4%; Score 6; DB 9; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRKKKK 144
|||||
DB 28 LRKKKK 33

RESULT 36
US-10-125-258-10
; Sequence 10, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
; APPLICANT: Altier, Daniel J.
; APPLICANT: Herrmann, Rafael
; APPLICANT: Lu, Albert L.
; APPLICANT: McCutchen, Billy F.
; APPLICANT: Presnail, James K.
; APPLICANT: Weaver, Janine L.
; APPLICANT: Wong, James F. H.
; TITLE OF INVENTION: Antimicrobial Polypeptides and Their
; FILE REFERENCE: 35718/246215
; CURRENT APPLICATION NUMBER: US/10/125,258
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/285,355
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Manduca sexta
; NAME/KEY: VARIANT
; LOCATION: 141
; OTHER INFORMATION: xaa = Any Amino Acid
US-10-125-258-10

Query Match 3.4%; Score 6; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AGLAYD 98
|||||
DB 86 AGLAYD 91

RESULT 37
US-09-983-802-196
; Sequence 196, Application US/09983802
; Publication No. US20030022185A1
; GENERAL INFORMATION:


```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41522
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALI61747.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.5
; OTHER INFORMATION: EST_HUMAN HIT: BF340331.1, EVALUE 2.00e-24
; OTHER INFORMATION: SWISSPROT HIT: P53236, EVALUE 1.50e+00
US-09-864-761-41522

```

```

Query Match          3.4%; Score 6; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 14 SLTLKL I9
    |||||
Db 19 SLTLKL 24

```

RESULT 30

```

US-10-092-154-844
; Sequence 844, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 844
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (54)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-844

```

```

Query Match          3.4%; Score 6; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 129 NLSVDV 134
    |||||
Db 69 NLSVDV 74

```

RESULT 31

```

US-09-764-847-844

```

```

; Sequence 844, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 844
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (54)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-844

```

```

Query Match          3.4%; Score 6; DB 10; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 129 NLSVDV 134
    |||||
Db 69 NLSVDV 74

```

RESULT 32

```

US-10-125-258-8
; Sequence 8, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
; APPLICANT: Altier, Daniel J.
; APPLICANT: Herrmann, Rafael
; APPLICANT: Lu, Albert L.
; APPLICANT: McCutchen, Billy F.
; APPLICANT: Presnail, James K.
; APPLICANT: Weaver, Janine L.
; APPLICANT: Wong, James F. H.
; TITLE OF INVENTION: Antimicrobial Polypeptides and Their
; FILE REFERENCE: 35718/246215
; CURRENT APPLICATION NUMBER: US/10/125,258
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/285,355
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Manduca sexta
US-10-125-258-8

```

```

Query Match          3.4%; Score 6; DB 9; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 93 AGLAYD 98
    |||||
Db 88 AGLAYD 93

```

RESULT 33

```

US-09-738-626-4318
; Sequence 4318, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43033
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049575.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: P45190, EVALUE 3.10e+00
; OTHER INFORMATION: EST_HUMAN HIT: AW904936.1, EVALUE 2.00e-31
US-09-864-761-43033

Query Match 3.4%; Score 6; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 KGLLLV 158
|||||
DB 31 KGLLLV 36

RESULT 27
US-09-867-550-612
; Sequence 612, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 612
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: wherein Xaa may be any one of Ala or Arg or Glu or Gln or Gly or
; OTHER INFORMATION: Leu or Lys or Pro or Ser or Thr or Val
US-09-867-550-612

Query Match 3.4%; Score 6; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LLLVEA 160
|||||
DB 26 LLLVEA 31

RESULT 28

US-10-300-616-49
; Sequence 49, Application US/10300616
; Publication No. US20030082801A1
; GENERAL INFORMATION:
; APPLICANT: BARNES, ASHLEY A.
; APPLICANT: WISE, ALAN
; APPLICANT: MARSHALL, FIONA H.
; APPLICANT: FRASER, NEIL J.
; APPLICANT: WHITE, JULIE H. M.
; APPLICANT: FOORD, STEVEN M.
; TITLE OF INVENTION: NOVEL RECEPTOR
; FILE REFERENCE: PG3558US2
; CURRENT APPLICATION NUMBER: US/10/300,616
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: GB9819420.2
; PRIOR FILING DATE: 1998-09-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-300-616-49

Query Match 3.4%; Score 6; DB 9; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 FTKKEE 70
|||||
DB 21 FTKKEE 26

RESULT 29
US-09-864-761-41522
; Sequence 41522, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods for Their Use
FILE REFERENCE: 11000.101c4u
CURRENT APPLICATION NUMBER: US/09/866.050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134
LENGTH: 90
TYPE: PRT
ORGANISM: Rat
US-09-866-050A-134

Query Match 3.4%; Score 6; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 99 KAASKT 104
| | | | |
Db 31 KAASKT 36

RESULT 25
US-09-864-761-43578
Sequence 43578 Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomlca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 43578
LENGTH: 90
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004147.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
OTHER INFORMATION: SWISSPROT HIT: Q62962, EVALUATE 6.00e-35
OTHER INFORMATION: EST_HUMAN HIT: A1277385.1, EVALUATE 2.00e-16
US-09-864-761-43578

Query Match 3.4%; Score 6; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 1; Caps 0;

QY 154 GLLLVE 159
| | | | |
Db 56 GLLLVE 61

RESULT 26
US-09-864-761-43033
Sequence 43033 Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomlca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661

US-10-079-854-110
; Sequence 110, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 76

; TYPE: PRT
; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature
; LOCATION: (17)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: misc_feature
; LOCATION: (66)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-079-854-110

Query Match 3.4%; Score 6; DB 9; Length 76;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LTLKLP 20
|||||
Db 53 LTLKLP 58

RESULT 21

US-09-764-878-110

; Sequence 110, Application US/09764878

; Patent No. US20020090615A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA121

; CURRENT APPLICATION NUMBER: US/09/764,878

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 428

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 110

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (17)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (66)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-878-110

Query Match 3.4%; Score 6; DB 10; Length 76;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LTLKLP 20
|||||
Db 53 LTLKLP 58

RESULT 22

US-09-893-737-12

; Sequence 12, Application US/09893737

; Patent No. US20020110855A1

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Presnell, Scott R.

; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS

; FILE REFERENCE: 00-41

; CURRENT APPLICATION NUMBER: US/09/893,737

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 60/215,446

; PRIOR FILING DATE: 2000-06-30

; NUMBER OF SEQ ID NOS: 329

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 12

; LENGTH: 80

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-893-737-12

Query Match 3.4%; Score 6; DB 10; Length 80;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 153 KGLLLV 158
|||||
Db 5 KGLLLV 10

RESULT 23

US-09-738-626-6220

; Sequence 6220, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 6220

; LENGTH: 82

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-6220

Query Match 3.4%; Score 6; DB 9; Length 82;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 DALTLR 92
|||||
Db 59 DALTLR 64

RESULT 24

US-09-866-050A-134

; Sequence 134, Application US/09866050A

; Publication No. US20030040471A1

; GENERAL INFORMATION:

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL36527.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.2
; OTHER INFORMATION: SWISSPROT HIT: P07954, EVALUE 1.00e-29
; OTHER INFORMATION: EST_HUMAN HIT: BE293788.1, EVALUE 1.00e-28
US-09-864-761-41976

Query Match 3.4%; Score 6; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 AYDKAA 101
| | | | |
Db 32 AYDKAA 37

RESULT 18
US-09-864-761-46121
; Sequence 46121, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
```

```
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46121
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010618.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.9
; OTHER INFORMATION: EST_HUMAN HIT: AW630291.1, EVALUE 9.00e-34
; OTHER INFORMATION: SWISSPROT HIT: Q03018, EVALUE 5.10e-01
US-09-864-761-46121
```

```
Query Match 3.4%; Score 6; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 157 LVEADY 162
| | | | |
Db 22 LVEADY 27
```

RESULT 19

```
US-09-925-300-1548
; Sequence 1548, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1548
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (37)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1548
```

```
Query Match 3.4%; Score 6; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 102 SKTHLS 107
| | | | |
Db 18 SKTHLS 23
```

RESULT 20

;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. I.1
;; SEQ ID NO 34896
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC010732.1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.5
;; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;; OTHER INFORMATION: EST_HUMAN HIT: BE541170.1, EVALUATE 8.80e-01

Query Match 3.4%; Score 6; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SUTLKL 19
Db 31 SUTLKL 36

RESULT 16
US-09-867-550-904
;; Sequence 904, Application US/09867550
;; Patent No. US20020082206A1
;; GENERAL INFORMATION:
;; APPLICANT: Leach, Martin D.
;; APPLICANT: Mehraban, Fuad,
;; APPLICANT: Conley, Pamela
;; APPLICANT: Law, Debbie
;; APPLICANT: Topper, James
;; TITLE OF INVENTION: No. US20020082206A1e1 Polynucleotides from Atherogenic Cells and
;; FILE REFERENCE: 21402-013 (Cura-313)
;; CURRENT APPLICATION NUMBER: US/09/867,550
;; CURRENT FILING DATE: 2001-09-20

;; PRIOR APPLICATION NUMBER: USSN 60/208,427
;; PRIOR FILING DATE: 2000-05-30
;; NUMBER OF SEQ ID NOS: 2125
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 904
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; US-09-867-550-904

Query Match 3.4%; Score 6; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 SKTHLS 107
Db 23 SKTHLS 28

RESULT 17
US-09-864-761-41976
;; Sequence 41976, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aemica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 41976
;; LENGTH: 68

```
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40107
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005317.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
; OTHER INFORMATION: EST_HUMAN HIT: BE615382.1, EVALUE 7.40e+00
; OTHER INFORMATION: SWISSPROT HIT: P07646, EVALUE 2.80e+00
US-09-864-761-40107

Query Match      3.4%; Score 6; DB 10; Length 37;
Best Local Similarity 100.0%; Pred.No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 LTLAG 94
Db 11 LTLAG 16

RESULT 13
US-10-122-822-25
; Sequence 25, Application US/10122822
; Publication No. US20030084477A1
; GENERAL INFORMATION:
; APPLICANT: Purdue Research Foundation
; TITLE OF INVENTION: GENES ENHANCING DISEASE RESISTANCE IN PLANTS
; FILE REFERENCE: 7024-511
; CURRENT APPLICATION NUMBER: US/10/122,822
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 09/202,161
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: PCT/US97/10382
; PRIOR FILING DATE: 1997-06-12
; PRIOR APPLICATION NUMBER: 60/046,494
; PRIOR FILING DATE: 1997-05-14
; PRIOR APPLICATION NUMBER: 60/019,633
; PRIOR FILING DATE: 1996-06-12
```

```
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: ASCII
; SEQ ID NO 25
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-122-822-25

Query Match      3.4%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred.No. 1.15e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 AYDKAA 101
Db 41 AYDKAA 46

RESULT 14
US-09-867-550-1856
; Sequence 1856, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1856
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1856

Query Match      3.4%; Score 6; DB 10; Length 60;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 154 GLLLE 159
Db 38 GLLLE 43

RESULT 15
US-09-864-761-34896
; Sequence 34896, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Acomica-X-I
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
```

;; PRIOR APPLICATION NUMBER: 60/070,923
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,007
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,057
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,006
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,369
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,367
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,169
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,008
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,365
;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 537
;; LENGTH: 24
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-097-065-537

Query Match 3.4%; Score 6; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SLTLKL 19
DB 12 SLTLKL 17
|||||

RESULT 11
US-09-864-761-36829
; Sequence 36829, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 36829
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC011174.2
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
US-09-864-761-36829

Query Match 3.4%; Score 6; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 KAASKT 104
DB 10 KAASKT 15
|||||

RESULT 12
US-09-864-761-40107
; Sequence 40107, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27

; LENGTH: 511
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-110

Query Match 3.9%; Score 7; DB 10; Length 511;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 KGSITLK 18
|||||
DB 283 KGSITLK 289

RESULT 7

US-09-738-626-3810
; Sequence 3810, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3810
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3810

Query Match 3.9%; Score 7; DB 9; Length 795;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 LLLVEAD 161
|||||
DB 141 LLLVEAD 147

RESULT 8

US-10-024-918-10
; Sequence 10, Application US/10024918
; Patent No. US20020168718A1
; GENERAL INFORMATION:
; APPLICANT: Hubbell, Jeffrey
; APPLICANT: Schense, Jason
; APPLICANT: Zisch, Andreas
; APPLICANT: Hall, Heike
; TITLE OF INVENTION: ENZYME-MEDIATED MODIFICATION OF FIBRIN FOR TISSUE ENGINEERING
; FILE REFERENCE: CIT 2606 CIP
; CURRENT APPLICATION NUMBER: US/10/024,918
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 8

; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: random coupling polylysine peptide attached to active peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: dansylated leucine
US-10-024-918-10

Query Match 3.4%; Score 6; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 LRKKKK 144
|||||
DB 1 LRKKKK 6

RESULT 9

US-10-150-111-181
; Sequence 181, Application US/10150111
; Publication No. US20030078386A1
; GENERAL INFORMATION:
; APPLICANT: Rubin et al.
; TITLE OF INVENTION: Secreted Protein HPEAD48
; FILE REFERENCE: P2018PDI
; CURRENT APPLICATION NUMBER: US/10/150,111
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 181
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-150-111-181

Query Match 3.4%; Score 6; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 LNDALT 90
|||||
DB 2 LNDALT 7

RESULT 10

US-10-097-065-537
; Sequence 537, Application US/10097065
; Publication No. US2003005236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17

APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6630
LENGTH: 93
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6630

Query Match 3.9%; Score 7; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 IKGLLLV 158
|||||||
DB 82 IKGLLLV 88

RESULT 4

US-10-278-173-150
Sequence 150, Application US/10278173
Publication No. US20030061637A1
GENERAL INFORMATION:
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Pineda, Onaira
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Pilgrim, Marsha
APPLICANT: Keddle, James
APPLICANT: Heard, Jacqueline
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION
FILE REFERENCE: MBI-009
CURRENT APPLICATION NUMBER: US/10/278,173
CURRENT FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US/09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 150
LENGTH: 218
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G26
US-10-278-173-150

Query Match 3.9%; Score 7; DB 9; Length 218;
Best Local Similarity 100.0%; Pred. No. 49;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 95 LAYDKAA 101
|||||||
DB 107 LAYDKAA 113

RESULT 5

US-10-278-173-82
Sequence 82, Application US/10278173
Publication No. US20030061637A1
GENERAL INFORMATION:
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Pineda, Onaira
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Pilgrim, Marsha
APPLICANT: Keddle, James
APPLICANT: Heard, Jacqueline
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION
FILE REFERENCE: MBI-009
CURRENT APPLICATION NUMBER: US/10/278,173
CURRENT FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US/09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 82
LENGTH: 334
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G5
US-10-278-173-82

Query Match 3.9%; Score 7; DB 9; Length 334;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 LAYDKAA 101
|||||||
DB 189 LAYDKAA 195

RESULT 6

US-09-881-752A-110
Sequence 110, Application US/09881752A
Patent No. US20020115078A1
GENERAL INFORMATION:
APPLICANT: Kleanthous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Millier, Charles
APPLICANT: Tomb, Jean-Francois
APPLICANT: Oomen, Raymond P.
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in t
FILE REFERENCE: Genome
FILE REFERENCE: 06132/041002
CURRENT APPLICATION NUMBER: US/09/881,752A
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 08/833,457
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 370
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 110

```
969 6 3.4 620 9 US-10-127-901A-324 Sequence 324, App
970 6 3.4 620 9 US-10-128-693A-324 Sequence 324, App
971 6 3.4 620 9 US-10-131-813A-324 Sequence 324, App
972 6 3.4 620 9 US-10-131-818A-324 Sequence 324, App
973 6 3.4 620 9 US-10-131-823A-324 Sequence 324, App
974 6 3.4 620 9 US-10-131-824A-324 Sequence 324, App
975 6 3.4 620 9 US-10-131-830A-324 Sequence 324, App
976 6 3.4 620 9 US-10-131-837A-324 Sequence 324, App
977 6 3.4 620 9 US-10-137-872A-324 Sequence 324, App
978 6 3.4 620 9 US-10-140-860-324 Sequence 324, App
979 6 3.4 620 9 US-10-142-417-324 Sequence 324, App
980 6 3.4 620 9 US-10-147-500-324 Sequence 324, App
981 6 3.4 620 9 US-10-147-502-324 Sequence 324, App
982 6 3.4 620 9 US-10-147-515-324 Sequence 324, App
983 6 3.4 620 9 US-10-147-517-324 Sequence 324, App
984 6 3.4 620 9 US-10-147-519-324 Sequence 324, App
985 6 3.4 620 9 US-10-147-526-324 Sequence 324, App
986 6 3.4 620 9 US-10-147-527-324 Sequence 324, App
987 6 3.4 620 9 US-10-152-395-324 Sequence 324, App
988 6 3.4 620 9 US-10-157-782-324 Sequence 324, App
989 6 3.4 620 9 US-09-902-634-73 Sequence 73, Appl
990 6 3.4 620 9 US-09-902-713-73 Sequence 73, Appl
991 6 3.4 620 9 US-09-907-979-73 Sequence 73, Appl
992 6 3.4 620 9 US-10-121-040-324 Sequence 324, App
993 6 3.4 620 9 US-10-121-056-324 Sequence 324, App
994 6 3.4 620 9 US-10-121-061-324 Sequence 324, App
995 6 3.4 620 9 US-10-123-235-324 Sequence 324, App
996 6 3.4 620 9 US-10-124-818-324 Sequence 324, App
997 6 3.4 620 9 US-10-125-926A-324 Sequence 324, App
998 6 3.4 620 9 US-10-125-930A-324 Sequence 324, App
999 6 3.4 620 9 US-10-127-831A-324 Sequence 324, App
1000 6 3.4 620 9 US-10-127-837A-324 Sequence 324, App
```

ALIGNMENTS

```
RESULT 1
US-09-996-140-11
; Sequence 11, Application US/09996140
; Patent No. US20020157136A1
; GENERAL INFORMATION:
; APPLICANT: Thomashow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zarka, Daniel
; TITLE OF INVENTION: PLANT HAVING ALTERED ENVIRONMENTAL
; FILE REFERENCE: MB10040US
; CURRENT APPLICATION NUMBER: US/09/996.140
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/706,270
; PRIOR FILING DATE: 1996-09-04
; PRIOR APPLICATION NUMBER: 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,235
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/198,119
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: PCT/US99/01895
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 11
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Tobacco
US-09-996-140-11
Query Match 3.9%; Score 7; DB 9; Length 61;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 95 LAYDKAA 101
Db 40 LAYDKAA 46
RESULT 2
US-09-996-140-134
; Sequence 134, Application US/09996140
; Patent No. US20020157136A1
; GENERAL INFORMATION:
; APPLICANT: Thomashow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zarka, Daniel
; APPLICANT: Jiang, Cai-Zhong
; TITLE OF INVENTION: PLANT HAVING ALTERED ENVIRONMENTAL
; FILE REFERENCE: MB10040US
; CURRENT APPLICATION NUMBER: US/09/996.140
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/706,270
; PRIOR FILING DATE: 1996-09-04
; PRIOR APPLICATION NUMBER: 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,235
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,575
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,227
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,234
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/198,119
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: PCT/US99/01895
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Tobacco
US-09-996-140-134
Query Match 3.9%; Score 7; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 95 LAYDKAA 101
Db 42 LAYDKAA 48
RESULT 3
US-09-738-626-6630
; Sequence 6630, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
```

823	6	3.4	458	9	US-10-207-919-496	Sequence 496, App	896	6	3.4	620	9	US-10-140-470-324	Sequence 324, App
824	6	3.4	458	9	US-10-207-920-496	Sequence 496, App	897	6	3.4	620	9	US-09-906-838-73	Sequence 73, Appl
825	6	3.4	458	9	US-10-207-925-496	Sequence 496, App	898	6	3.4	620	9	US-09-907-613-73	Sequence 73, Appl
826	6	3.4	458	9	US-10-208-021-496	Sequence 496, App	899	6	3.4	620	9	US-09-907-942-73	Sequence 73, Appl
827	6	3.4	458	9	US-10-208-022-496	Sequence 496, App	900	6	3.4	620	9	US-10-175-746-324	Sequence 324, App
828	6	3.4	458	9	US-10-208-023-496	Sequence 496, App	901	6	3.4	620	9	US-10-176-918-324	Sequence 324, App
829	6	3.4	458	9	US-10-208-026-496	Sequence 496, App	902	6	3.4	620	9	US-10-176-921-324	Sequence 324, App
830	6	3.4	458	9	US-10-208-029-496	Sequence 496, App	903	6	3.4	620	9	US-10-137-865-324	Sequence 324, App
831	6	3.4	458	9	US-10-208-030-496	Sequence 496, App	904	6	3.4	620	9	US-10-137-865-324	Sequence 324, App
832	6	3.4	458	9	US-10-232-232-496	Sequence 496, App	905	6	3.4	620	9	US-10-140-470-324	Sequence 324, App
833	6	3.4	458	9	US-10-173-693-496	Sequence 496, App	906	6	3.4	620	9	US-09-904-820-73	Sequence 73, Appl
834	6	3.4	458	9	US-10-174-578-496	Sequence 496, App	907	6	3.4	620	9	US-09-904-859-73	Sequence 73, Appl
835	6	3.4	458	9	US-10-175-741-496	Sequence 496, App	908	6	3.4	620	9	US-09-909-204-73	Sequence 73, Appl
836	6	3.4	458	9	US-10-175-750-496	Sequence 496, App	909	6	3.4	620	9	US-10-142-431-324	Sequence 324, App
837	6	3.4	458	9	US-10-176-986-496	Sequence 496, App	910	6	3.4	620	9	US-10-143-114-324	Sequence 324, App
838	6	3.4	458	9	US-10-184-641-496	Sequence 496, App	911	6	3.4	620	9	US-09-904-786-73	Sequence 73, Appl
839	6	3.4	458	9	US-10-187-888-496	Sequence 496, App	912	6	3.4	620	9	US-09-906-646-73	Sequence 73, Appl
840	6	3.4	458	9	US-10-194-360-496	Sequence 496, App	913	6	3.4	620	9	US-09-906-700-73	Sequence 73, Appl
841	6	3.4	458	9	US-10-194-365-496	Sequence 496, App	914	6	3.4	620	9	US-10-140-002-324	Sequence 324, App
842	6	3.4	458	9	US-10-195-895-496	Sequence 496, App	915	6	3.4	620	9	US-09-902-903-73	Sequence 73, Appl
843	6	3.4	458	9	US-10-195-898-496	Sequence 496, App	916	6	3.4	620	9	US-09-903-749A-73	Sequence 73, Appl
844	6	3.4	458	9	US-10-196-759-496	Sequence 496, App	917	6	3.4	620	9	US-09-903-786-73	Sequence 73, Appl
845	6	3.4	458	9	US-10-199-302-496	Sequence 496, App	918	6	3.4	620	9	US-10-142-419-324	Sequence 324, App
846	6	3.4	458	9	US-10-201-323-496	Sequence 496, App	919	6	3.4	620	9	US-09-902-738-73	Sequence 73, Appl
847	6	3.4	458	9	US-10-205-510-496	Sequence 496, App	920	6	3.4	620	9	US-09-904-119-73	Sequence 73, Appl
848	6	3.4	458	9	US-10-205-891-496	Sequence 496, App	921	6	3.4	620	9	US-09-904-956-73	Sequence 73, Appl
849	6	3.4	458	9	US-10-205-904-496	Sequence 496, App	922	6	3.4	620	9	US-09-907-794-73	Sequence 73, Appl
850	6	3.4	458	9	US-10-206-917-496	Sequence 496, App	923	6	3.4	620	9	US-10-123-263-324	Sequence 324, App
851	6	3.4	458	9	US-10-207-923-496	Sequence 496, App	924	6	3.4	620	9	US-10-142-423-324	Sequence 324, App
852	6	3.4	458	9	US-10-207-924-496	Sequence 496, App	925	6	3.4	620	9	US-09-902-692-73	Sequence 73, Appl
853	6	3.4	458	9	US-10-208-028-496	Sequence 496, App	926	6	3.4	620	9	US-09-903-520-73	Sequence 73, Appl
854	6	3.4	458	9	US-10-121-062-496	Sequence 496, App	927	6	3.4	620	9	US-09-903-943-73	Sequence 73, Appl
855	6	3.4	458	9	US-10-175-753-496	Sequence 496, App	928	6	3.4	620	9	US-09-904-462-73	Sequence 73, Appl
856	6	3.4	458	9	US-10-180-553-496	Sequence 496, App	929	6	3.4	620	9	US-09-905-056-73	Sequence 73, Appl
857	6	3.4	458	9	US-10-201-327-496	Sequence 496, App	930	6	3.4	620	9	US-09-907-925-73	Sequence 73, Appl
858	6	3.4	458	9	US-10-173-696-496	Sequence 496, App	931	6	3.4	620	9	US-10-121-050-324	Sequence 324, App
859	6	3.4	458	9	US-10-183-003-496	Sequence 496, App	932	6	3.4	620	9	US-10-141-753-324	Sequence 324, App
860	6	3.4	458	9	US-10-183-016-496	Sequence 496, App	933	6	3.4	620	9	US-09-904-553-73	Sequence 73, Appl
861	6	3.4	458	12	US-10-052-586-496	Sequence 496, App	934	6	3.4	620	9	US-09-905-381-73	Sequence 73, Appl
862	6	3.4	476	12	US-10-007-693-69	Sequence 69, Appl	935	6	3.4	620	9	US-09-909-064-73	Sequence 73, Appl
863	6	3.4	478	9	US-10-105-717-2	Sequence 2, Appl	936	6	3.4	620	9	US-10-143-032-324	Sequence 324, App
864	6	3.4	481	10	US-09-158-180-2	Sequence 2, Appl	937	6	3.4	620	9	US-10-123-108-324	Sequence 324, App
865	6	3.4	495	10	US-09-823-356-4	Sequence 4, Appl	938	6	3.4	620	9	US-10-123-236-324	Sequence 324, App
866	6	3.4	496	10	US-09-815-242-13401	Sequence 13401, A	939	6	3.4	620	9	US-10-123-261-324	Sequence 324, App
867	6	3.4	502	10	US-09-815-242-5441	Sequence 5441, Ap	940	6	3.4	620	9	US-10-140-921-324	Sequence 324, App
868	6	3.4	502	10	US-09-815-242-12387	Sequence 12387, A	941	6	3.4	620	9	US-10-140-928-324	Sequence 324, App
869	6	3.4	504	10	US-09-815-242-12821	Sequence 12821, A	942	6	3.4	620	9	US-09-905-088-73	Sequence 73, Appl
870	6	3.4	504	9	US-09-972-268-8	Sequence 8, Appl	943	6	3.4	620	9	US-09-907-575-73	Sequence 73, Appl
871	6	3.4	510	9	US-09-959-845-4	Sequence 4, Appl	944	6	3.4	620	9	US-10-121-045-324	Sequence 324, App
872	6	3.4	510	9	US-09-972-268-10	Sequence 10, Appl	945	6	3.4	620	9	US-10-123-292-324	Sequence 324, App
873	6	3.4	510	9	US-09-972-268-12	Sequence 12, Appl	946	6	3.4	620	9	US-10-123-903-324	Sequence 324, App
874	6	3.4	510	9	US-09-972-268-18	Sequence 18, Appl	947	6	3.4	620	9	US-10-124-819-324	Sequence 324, App
875	6	3.4	512	9	US-09-983-204-14	Sequence 14, Appl	948	6	3.4	620	9	US-10-124-822-324	Sequence 324, App
876	6	3.4	542	9	US-09-972-268-2	Sequence 2, Appl	949	6	3.4	620	9	US-10-140-925-324	Sequence 324, App
877	6	3.4	549	9	US-09-959-845-2	Sequence 2, Appl	950	6	3.4	620	9	US-10-160-498-324	Sequence 324, App
878	6	3.4	549	9	US-09-972-268-4	Sequence 4, Appl	951	6	3.4	620	9	US-09-902-759-73	Sequence 73, Appl
879	6	3.4	549	9	US-09-972-268-6	Sequence 6, Appl	952	6	3.4	620	9	US-09-905-075-73	Sequence 73, Appl
880	6	3.4	549	9	US-09-972-268-17	Sequence 17, Appl	953	6	3.4	620	9	US-10-121-041-324	Sequence 324, App
881	6	3.4	569	9	US-09-906-419-39	Sequence 39, Appl	954	6	3.4	620	9	US-10-121-043-324	Sequence 324, App
882	6	3.4	594	9	US-10-255-969-10	Sequence 10, Appl	955	6	3.4	620	9	US-10-121-047-324	Sequence 324, App
883	6	3.4	594	12	US-10-139-262-10	Sequence 10, Appl	956	6	3.4	620	9	US-10-123-215-324	Sequence 324, App
884	6	3.4	595	9	US-09-972-268-14	Sequence 14, Appl	957	6	3.4	620	9	US-10-123-902-324	Sequence 324, App
885	6	3.4	614	10	US-09-803-589-4	Sequence 4, Appl	958	6	3.4	620	9	US-10-123-908-324	Sequence 324, App
886	6	3.4	620	9	US-09-905-291A-73	Sequence 73, Appl	959	6	3.4	620	9	US-10-123-909-324	Sequence 324, App
887	6	3.4	620	9	US-09-965-529-11	Sequence 11, Appl	960	6	3.4	620	9	US-10-123-910-324	Sequence 324, App
888	6	3.4	620	9	US-09-902-853-73	Sequence 73, Appl	961	6	3.4	620	9	US-10-124-813-324	Sequence 324, App
889	6	3.4	620	9	US-09-907-824-73	Sequence 73, Appl	962	6	3.4	620	9	US-10-124-817-324	Sequence 324, App
890	6	3.4	620	9	US-09-907-841-73	Sequence 73, Appl	963	6	3.4	620	9	US-10-125-922-324	Sequence 324, App
891	6	3.4	620	9	US-09-904-011-73	Sequence 73, Appl	964	6	3.4	620	9	US-10-125-924-324	Sequence 324, App
892	6	3.4	620	9	US-10-028-072-324	Sequence 324, App	965	6	3.4	620	9	US-10-127-825A-324	Sequence 324, App
893	6	3.4	620	9	US-09-906-742-73	Sequence 73, Appl	966	6	3.4	620	9	US-10-127-829A-324	Sequence 324, App
894	6	3.4	620	9	US-10-121-049-324	Sequence 324, App	967	6	3.4	620	9	US-10-127-835A-324	Sequence 324, App
895	6	3.4	620	9	US-10-123-904-324	Sequence 324, App	968	6	3.4	620	9	US-10-127-839A-324	Sequence 324, App

677	6	3.4	458	9	US-10-199-314-496	Sequence 496, App	750	9	3.4	458	9	US-10-192-014-496	Sequence 496, App
678	6	3.4	458	9	US-10-199-317-496	Sequence 496, App	751	6	3.4	458	9	US-10-192-016-496	Sequence 496, App
679	6	3.4	458	9	US-10-199-665-496	Sequence 496, App	752	6	3.4	458	9	US-10-194-362-496	Sequence 496, App
680	6	3.4	458	9	US-10-199-666-496	Sequence 496, App	753	6	3.4	458	9	US-10-194-364-496	Sequence 496, App
681	6	3.4	458	9	US-10-199-669-496	Sequence 496, App	754	6	3.4	458	9	US-10-194-395-496	Sequence 496, App
682	6	3.4	458	9	US-10-201-534-496	Sequence 496, App	755	6	3.4	458	9	US-10-194-424-496	Sequence 496, App
683	6	3.4	458	9	US-10-201-770-496	Sequence 496, App	756	6	3.4	458	9	US-10-194-458-496	Sequence 496, App
684	6	3.4	458	9	US-10-201-853-496	Sequence 496, App	757	6	3.4	458	9	US-10-194-459-496	Sequence 496, App
685	6	3.4	458	9	US-10-201-856-496	Sequence 496, App	758	6	3.4	458	9	US-10-194-488-496	Sequence 496, App
686	6	3.4	458	9	US-10-202-469-496	Sequence 496, App	759	6	3.4	458	9	US-10-195-886-496	Sequence 496, App
687	6	3.4	458	9	US-10-202-470-496	Sequence 496, App	760	6	3.4	458	9	US-10-195-891-496	Sequence 496, App
688	6	3.4	458	9	US-10-202-476-496	Sequence 496, App	761	6	3.4	458	9	US-10-196-746-496	Sequence 496, App
689	6	3.4	458	9	US-10-202-934-496	Sequence 496, App	762	6	3.4	458	9	US-10-196-752-496	Sequence 496, App
690	6	3.4	458	9	US-10-202-933-496	Sequence 496, App	763	6	3.4	458	9	US-10-196-753-496	Sequence 496, App
691	6	3.4	458	9	US-10-202-936-496	Sequence 496, App	764	6	3.4	458	9	US-10-196-754-496	Sequence 496, App
692	6	3.4	458	9	US-10-202-939-496	Sequence 496, App	765	6	3.4	458	9	US-10-196-761-496	Sequence 496, App
693	6	3.4	458	9	US-10-205-504-496	Sequence 496, App	766	6	3.4	458	9	US-10-197-692-496	Sequence 496, App
694	6	3.4	458	9	US-10-205-504-496	Sequence 496, App	767	6	3.4	458	9	US-10-197-693-496	Sequence 496, App
695	6	3.4	458	9	US-10-205-509-496	Sequence 496, App	768	6	3.4	458	9	US-10-197-696-496	Sequence 496, App
696	6	3.4	458	9	US-10-205-893-496	Sequence 496, App	769	6	3.4	458	9	US-10-197-698-496	Sequence 496, App
697	6	3.4	458	9	US-10-205-899-496	Sequence 496, App	770	6	3.4	458	9	US-10-197-703-496	Sequence 496, App
698	6	3.4	458	9	US-10-205-900-496	Sequence 496, App	771	6	3.4	458	9	US-10-197-711-496	Sequence 496, App
699	6	3.4	458	9	US-10-205-909-496	Sequence 496, App	772	6	3.4	458	9	US-10-198-757-496	Sequence 496, App
700	6	3.4	458	9	US-10-183-002-496	Sequence 496, App	773	6	3.4	458	9	US-10-198-761-496	Sequence 496, App
701	6	3.4	458	9	US-10-184-621-496	Sequence 496, App	774	6	3.4	458	9	US-10-198-762-496	Sequence 496, App
702	6	3.4	458	9	US-10-184-638-496	Sequence 496, App	775	6	3.4	458	9	US-10-198-763-496	Sequence 496, App
703	6	3.4	458	9	US-10-187-752-496	Sequence 496, App	776	6	3.4	458	9	US-10-198-767-496	Sequence 496, App
704	6	3.4	458	9	US-10-187-887-496	Sequence 496, App	777	6	3.4	458	9	US-10-199-301-496	Sequence 496, App
705	6	3.4	458	9	US-10-194-461-496	Sequence 496, App	778	6	3.4	458	9	US-10-199-307-496	Sequence 496, App
706	6	3.4	458	9	US-10-195-892-496	Sequence 496, App	779	6	3.4	458	9	US-10-199-312-496	Sequence 496, App
707	6	3.4	458	9	US-10-196-751-496	Sequence 496, App	780	6	3.4	458	9	US-10-199-315-496	Sequence 496, App
708	6	3.4	458	9	US-10-197-697-496	Sequence 496, App	781	6	3.4	458	9	US-10-199-316-496	Sequence 496, App
709	6	3.4	458	9	US-10-197-707-496	Sequence 496, App	782	6	3.4	458	9	US-10-199-457-496	Sequence 496, App
710	6	3.4	458	9	US-10-199-303-496	Sequence 496, App	783	6	3.4	458	9	US-10-199-459-496	Sequence 496, App
711	6	3.4	458	9	US-10-199-318-496	Sequence 496, App	784	6	3.4	458	9	US-10-199-460-496	Sequence 496, App
712	6	3.4	458	9	US-10-199-458-496	Sequence 496, App	785	6	3.4	458	9	US-10-199-461-496	Sequence 496, App
713	6	3.4	458	9	US-10-199-462-496	Sequence 496, App	786	6	3.4	458	9	US-10-199-667-496	Sequence 496, App
714	6	3.4	458	9	US-10-201-328-496	Sequence 496, App	787	6	3.4	458	9	US-10-199-673-496	Sequence 496, App
715	6	3.4	458	9	US-10-201-328-496	Sequence 496, App	788	6	3.4	458	9	US-10-201-321-496	Sequence 496, App
716	6	3.4	458	9	US-10-201-527-496	Sequence 496, App	789	6	3.4	458	9	US-10-201-322-496	Sequence 496, App
717	6	3.4	458	9	US-10-201-528-496	Sequence 496, App	790	6	3.4	458	9	US-10-201-326-496	Sequence 496, App
718	6	3.4	458	9	US-10-201-529-496	Sequence 496, App	791	6	3.4	458	9	US-10-201-532-496	Sequence 496, App
719	6	3.4	458	9	US-10-201-530-496	Sequence 496, App	792	6	3.4	458	9	US-10-201-533-496	Sequence 496, App
720	6	3.4	458	9	US-10-202-408-496	Sequence 496, App	793	6	3.4	458	9	US-10-201-535-496	Sequence 496, App
721	6	3.4	458	9	US-10-202-409-496	Sequence 496, App	794	6	3.4	458	9	US-10-201-769-496	Sequence 496, App
722	6	3.4	458	9	US-10-202-411-496	Sequence 496, App	795	6	3.4	458	9	US-10-201-771-496	Sequence 496, App
723	6	3.4	458	9	US-10-202-472-496	Sequence 496, App	796	6	3.4	458	9	US-10-201-854-496	Sequence 496, App
724	6	3.4	458	9	US-10-205-502-496	Sequence 496, App	797	6	3.4	458	9	US-10-202-410-496	Sequence 496, App
725	6	3.4	458	9	US-10-205-507-496	Sequence 496, App	798	6	3.4	458	9	US-10-202-473-496	Sequence 496, App
726	6	3.4	458	9	US-10-205-511-496	Sequence 496, App	799	6	3.4	458	9	US-10-202-474-496	Sequence 496, App
727	6	3.4	458	9	US-10-205-903-496	Sequence 496, App	800	6	3.4	458	9	US-10-205-503-496	Sequence 496, App
728	6	3.4	458	9	US-10-205-907-496	Sequence 496, App	801	6	3.4	458	9	US-10-205-512-496	Sequence 496, App
729	6	3.4	458	9	US-10-194-456-496	Sequence 496, App	802	6	3.4	458	9	US-10-205-892-496	Sequence 496, App
730	6	3.4	458	9	US-10-196-758-496	Sequence 496, App	803	6	3.4	458	9	US-10-205-894-496	Sequence 496, App
731	6	3.4	458	9	US-10-198-770-496	Sequence 496, App	804	6	3.4	458	9	US-10-205-896-496	Sequence 496, App
732	6	3.4	458	9	US-10-199-308-496	Sequence 496, App	805	6	3.4	458	9	US-10-205-898-496	Sequence 496, App
733	6	3.4	458	9	US-10-200-617-496	Sequence 496, App	806	6	3.4	458	9	US-10-205-901-496	Sequence 496, App
734	6	3.4	458	9	US-10-205-893-496	Sequence 496, App	807	6	3.4	458	9	US-10-205-903-496	Sequence 496, App
735	6	3.4	458	9	US-10-205-897-496	Sequence 496, App	808	6	3.4	458	9	US-10-206-909-496	Sequence 496, App
736	6	3.4	458	9	US-10-174-571-496	Sequence 496, App	809	6	3.4	458	9	US-10-206-910-496	Sequence 496, App
737	6	3.4	458	9	US-10-176-746-496	Sequence 496, App	810	6	3.4	458	9	US-10-206-911-496	Sequence 496, App
738	6	3.4	458	9	US-10-176-923-496	Sequence 496, App	811	6	3.4	458	9	US-10-206-912-496	Sequence 496, App
739	6	3.4	458	9	US-10-183-011-496	Sequence 496, App	812	6	3.4	458	9	US-10-206-913-496	Sequence 496, App
740	6	3.4	458	9	US-10-184-633-496	Sequence 496, App	813	6	3.4	458	9	US-10-206-914-496	Sequence 496, App
741	6	3.4	458	9	US-10-184-639-496	Sequence 496, App	814	6	3.4	458	9	US-10-206-920-496	Sequence 496, App
742	6	3.4	458	9	US-10-187-742-496	Sequence 496, App	815	6	3.4	458	9	US-10-206-921-496	Sequence 496, App
743	6	3.4	458	9	US-10-187-748-496	Sequence 496, App	816	6	3.4	458	9	US-10-206-923-496	Sequence 496, App
744	6	3.4	458	9	US-10-188-766-496	Sequence 496, App	817	6	3.4	458	9	US-10-206-925-496	Sequence 496, App
745	6	3.4	458	9	US-10-188-771-496	Sequence 496, App	818	6	3.4	458	9	US-10-206-926-496	Sequence 496, App
746	6	3.4	458	9	US-10-192-006-496	Sequence 496, App	819	6	3.4	458	9	US-10-206-927-496	Sequence 496, App
747	6	3.4	458	9	US-10-192-008-496	Sequence 496, App	820	6	3.4	458	9	US-10-207-916-496	Sequence 496, App
748	6	3.4	458	9	US-10-192-009-496	Sequence 496, App	821	6	3.4	458	9	US-10-207-917-496	Sequence 496, App
749	6	3.4	458	9	US-10-192-012-496	Sequence 496, App	822	6	3.4	458	9	US-10-207-918-496	Sequence 496, App

531	6	3.4	458	9	US-10-180-543-496	Sequence 496, App	604	6	3.4	458	9	US-10-176-479-496	Sequence 496, App
532	6	3.4	458	9	US-10-180-544-496	Sequence 496, App	605	6	3.4	458	9	US-10-176-748-496	Sequence 496, App
533	6	3.4	458	9	US-10-180-546-496	Sequence 496, App	606	6	3.4	458	9	US-10-176-916-496	Sequence 496, App
534	6	3.4	458	9	US-10-180-547-496	Sequence 496, App	607	6	3.4	458	9	US-10-179-507-496	Sequence 496, App
535	6	3.4	458	9	US-10-180-549-496	Sequence 496, App	608	6	3.4	458	9	US-10-179-516-496	Sequence 496, App
536	6	3.4	458	9	US-10-180-555-496	Sequence 496, App	609	6	3.4	458	9	US-10-179-519-496	Sequence 496, App
537	6	3.4	458	9	US-10-180-553-496	Sequence 496, App	610	6	3.4	458	9	US-10-179-525-496	Sequence 496, App
538	6	3.4	458	9	US-10-181-000-496	Sequence 496, App	611	6	3.4	458	9	US-10-180-540-496	Sequence 496, App
539	6	3.4	458	9	US-10-183-010-496	Sequence 496, App	612	6	3.4	458	9	US-10-180-545-496	Sequence 496, App
540	6	3.4	458	9	US-10-183-013-496	Sequence 496, App	613	6	3.4	458	9	US-10-183-006-496	Sequence 496, App
541	6	3.4	458	9	US-10-184-614-496	Sequence 496, App	614	6	3.4	458	9	US-10-183-008-496	Sequence 496, App
542	6	3.4	458	9	US-10-184-623-496	Sequence 496, App	615	6	3.4	458	9	US-10-183-017-496	Sequence 496, App
543	6	3.4	458	9	US-10-184-633-496	Sequence 496, App	616	6	3.4	458	9	US-10-183-019-496	Sequence 496, App
544	6	3.4	458	9	US-10-184-637-496	Sequence 496, App	617	6	3.4	458	9	US-10-184-618-496	Sequence 496, App
545	6	3.4	458	9	US-10-184-646-496	Sequence 496, App	618	6	3.4	458	9	US-10-184-625-496	Sequence 496, App
546	6	3.4	458	9	US-10-184-647-496	Sequence 496, App	619	6	3.4	458	9	US-10-184-626-496	Sequence 496, App
547	6	3.4	458	9	US-10-184-652-496	Sequence 496, App	620	6	3.4	458	9	US-10-184-627-496	Sequence 496, App
548	6	3.4	458	9	US-10-187-594-496	Sequence 496, App	621	6	3.4	458	9	US-10-184-645-496	Sequence 496, App
549	6	3.4	458	9	US-10-187-596-496	Sequence 496, App	622	6	3.4	458	9	US-10-184-654-496	Sequence 496, App
550	6	3.4	458	9	US-10-187-743-496	Sequence 496, App	623	6	3.4	458	9	US-10-184-655-496	Sequence 496, App
551	6	3.4	458	9	US-10-187-885-496	Sequence 496, App	624	6	3.4	458	9	US-10-188-774-496	Sequence 496, App
552	6	3.4	458	9	US-10-187-886-496	Sequence 496, App	625	6	3.4	458	9	US-10-188-775-496	Sequence 496, App
553	6	3.4	458	9	US-10-199-464-496	Sequence 496, App	626	6	3.4	458	9	US-10-194-462-496	Sequence 496, App
554	6	3.4	458	9	US-10-176-751-496	Sequence 496, App	627	6	3.4	458	9	US-10-195-902-496	Sequence 496, App
555	6	3.4	458	9	US-10-176-760-496	Sequence 496, App	628	6	3.4	458	9	US-10-196-743-496	Sequence 496, App
556	6	3.4	458	9	US-10-176-990-496	Sequence 496, App	629	6	3.4	458	9	US-10-196-745-496	Sequence 496, App
557	6	3.4	458	9	US-10-180-541-496	Sequence 496, App	630	6	3.4	458	9	US-10-196-762-496	Sequence 496, App
558	6	3.4	458	9	US-10-180-542-496	Sequence 496, App	631	6	3.4	458	9	US-10-197-695-496	Sequence 496, App
559	6	3.4	458	9	US-10-180-548-496	Sequence 496, App	632	6	3.4	458	9	US-10-176-484-496	Sequence 496, App
560	6	3.4	458	9	US-10-180-553-496	Sequence 496, App	633	6	3.4	458	9	US-10-176-753-496	Sequence 496, App
561	6	3.4	458	9	US-10-180-998-496	Sequence 496, App	634	6	3.4	458	9	US-10-176-917-496	Sequence 496, App
562	6	3.4	458	9	US-10-180-999-496	Sequence 496, App	635	6	3.4	458	9	US-10-176-982-496	Sequence 496, App
563	6	3.4	458	9	US-10-183-013-496	Sequence 496, App	636	6	3.4	458	9	US-10-179-506-496	Sequence 496, App
564	6	3.4	458	9	US-10-184-612-496	Sequence 496, App	637	6	3.4	458	9	US-10-179-513-496	Sequence 496, App
565	6	3.4	458	9	US-10-184-615-496	Sequence 496, App	638	6	3.4	458	9	US-10-179-514-496	Sequence 496, App
566	6	3.4	458	9	US-10-184-617-496	Sequence 496, App	639	6	3.4	458	9	US-10-179-522-496	Sequence 496, App
567	6	3.4	458	9	US-10-184-622-496	Sequence 496, App	640	6	3.4	458	9	US-10-180-556-496	Sequence 496, App
568	6	3.4	458	9	US-10-184-628-496	Sequence 496, App	641	6	3.4	458	9	US-10-180-560-496	Sequence 496, App
569	6	3.4	458	9	US-10-184-629-496	Sequence 496, App	642	6	3.4	458	9	US-10-183-015-496	Sequence 496, App
570	6	3.4	458	9	US-10-184-630-496	Sequence 496, App	643	6	3.4	458	9	US-10-184-618-496	Sequence 496, App
571	6	3.4	458	9	US-10-184-631-496	Sequence 496, App	644	6	3.4	458	9	US-10-184-620-496	Sequence 496, App
572	6	3.4	458	9	US-10-184-632-496	Sequence 496, App	645	6	3.4	458	9	US-10-184-643-496	Sequence 496, App
573	6	3.4	458	9	US-10-184-634-496	Sequence 496, App	646	6	3.4	458	9	US-10-184-656-496	Sequence 496, App
574	6	3.4	458	9	US-10-184-640-496	Sequence 496, App	647	6	3.4	458	9	US-10-192-010-496	Sequence 496, App
575	6	3.4	458	9	US-10-184-650-496	Sequence 496, App	648	6	3.4	458	9	US-10-195-894-496	Sequence 496, App
576	6	3.4	458	9	US-10-184-651-496	Sequence 496, App	649	6	3.4	458	9	US-10-205-908-496	Sequence 496, App
577	6	3.4	458	9	US-10-187-588-496	Sequence 496, App	650	6	3.4	458	9	US-10-184-619-496	Sequence 496, App
578	6	3.4	458	9	US-10-187-597-496	Sequence 496, App	651	6	3.4	458	9	US-10-184-855-496	Sequence 496, App
579	6	3.4	458	9	US-10-187-598-496	Sequence 496, App	652	6	3.4	458	9	US-10-194-425-496	Sequence 496, App
580	6	3.4	458	9	US-10-187-600-496	Sequence 496, App	653	6	3.4	458	9	US-10-194-485-496	Sequence 496, App
581	6	3.4	458	9	US-10-187-601-496	Sequence 496, App	654	6	3.4	458	9	US-10-195-885-496	Sequence 496, App
582	6	3.4	458	9	US-10-187-602-496	Sequence 496, App	655	6	3.4	458	9	US-10-195-890-496	Sequence 496, App
583	6	3.4	458	9	US-10-187-603-496	Sequence 496, App	656	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
584	6	3.4	458	9	US-10-187-741-496	Sequence 496, App	657	6	3.4	458	9	US-10-194-394-496	Sequence 496, App
585	6	3.4	458	9	US-10-187-743-496	Sequence 496, App	658	6	3.4	458	9	US-10-194-425-496	Sequence 496, App
586	6	3.4	458	9	US-10-187-746-496	Sequence 496, App	659	6	3.4	458	9	US-10-187-599-496	Sequence 496, App
587	6	3.4	458	9	US-10-187-747-496	Sequence 496, App	660	6	3.4	458	9	US-10-187-750-496	Sequence 496, App
588	6	3.4	458	9	US-10-187-751-496	Sequence 496, App	661	6	3.4	458	9	US-10-188-780-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-192-015-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	458	9	US-10-187-753-496	Sequence 496, App	662	6	3.4	458	9	US-10-195-899-496	Sequence 496, App
589	6	3.4	45										

385	6	3.4	359	9	US-10-199-673-42	Sequence 42, Appl	458	6	3.4	359	9	US-10-180-553-12	Sequence 42, Appl
386	6	3.4	359	9	US-10-201-321-42	Sequence 42, Appl	459	6	3.4	359	9	US-10-201-327-42	Sequence 42, Appl
387	6	3.4	359	9	US-10-201-322-42	Sequence 42, Appl	460	6	3.4	359	9	US-09-978-757A-45	Sequence 45, Appl
388	6	3.4	359	9	US-10-201-326-42	Sequence 42, Appl	461	6	3.4	359	9	US-10-173-696-42	Sequence 42, Appl
389	6	3.4	359	9	US-10-201-532-42	Sequence 42, Appl	462	6	3.4	359	9	US-10-183-003-42	Sequence 42, Appl
390	6	3.4	359	9	US-10-201-533-42	Sequence 42, Appl	463	6	3.4	359	9	US-10-183-016-42	Sequence 42, Appl
391	6	3.4	359	9	US-10-201-535-42	Sequence 42, Appl	464	6	3.4	359	12	US-10-052-586-42	Sequence 6, Appl
392	6	3.4	359	9	US-10-201-769-42	Sequence 42, Appl	465	6	3.4	359	9	US-10-060-218A-6	Sequence 68, Appl
393	6	3.4	359	9	US-10-201-771-42	Sequence 42, Appl	466	6	3.4	374	9	US-09-847-208-68	Sequence 16, Appl
394	6	3.4	359	9	US-10-201-854-42	Sequence 42, Appl	467	6	3.4	387	9	US-09-972-268-16	Sequence 12, Appl
395	6	3.4	359	9	US-10-202-410-42	Sequence 42, Appl	468	6	3.4	393	9	US-09-977-418-12	Sequence 12, Appl
396	6	3.4	359	9	US-10-202-473-42	Sequence 42, Appl	469	6	3.4	393	9	US-09-977-033A-12	Sequence 12, Appl
397	6	3.4	359	9	US-10-202-474-42	Sequence 42, Appl	470	6	3.4	393	10	US-09-840-787-11	Sequence 11, Appl
398	6	3.4	359	9	US-10-205-503-42	Sequence 42, Appl	471	6	3.4	394	9	US-09-738-626-4096	Sequence 4096, A
399	6	3.4	359	9	US-10-205-512-42	Sequence 42, Appl	472	6	3.4	395	10	US-09-815-242-13892	Sequence 13892, A
400	6	3.4	359	9	US-10-205-892-42	Sequence 42, Appl	473	6	3.4	401	10	US-09-815-242-13828	Sequence 13828, A
401	6	3.4	359	9	US-10-205-894-42	Sequence 42, Appl	474	6	3.4	410	9	US-09-910-186A-32	Sequence 32, Appl
402	6	3.4	359	9	US-10-205-896-42	Sequence 42, Appl	475	6	3.4	411	9	US-09-866-050A-302	Sequence 302, App
403	6	3.4	359	9	US-10-205-898-42	Sequence 42, Appl	476	6	3.4	421	9	US-10-012-055-2	Sequence 2, Appl
404	6	3.4	359	9	US-10-205-901-42	Sequence 42, Appl	477	6	3.4	426	9	US-09-972-268-15	Sequence 15, Appl
405	6	3.4	359	9	US-10-205-903-42	Sequence 42, Appl	478	6	3.4	434	10	US-09-815-242-11827	Sequence 11827, A
406	6	3.4	359	9	US-10-206-909-42	Sequence 42, Appl	479	6	3.4	437	9	US-09-972-268-31	Sequence 31, Appl
407	6	3.4	359	9	US-10-206-910-42	Sequence 42, Appl	480	6	3.4	437	10	US-09-797-207-20	Sequence 20, Appl
408	6	3.4	359	9	US-10-206-911-42	Sequence 42, Appl	481	6	3.4	438	9	US-09-959-845-6	Sequence 6, Appl
409	6	3.4	359	9	US-10-206-912-42	Sequence 42, Appl	482	6	3.4	438	9	US-09-972-268-19	Sequence 19, Appl
410	6	3.4	359	9	US-10-206-913-42	Sequence 42, Appl	483	6	3.4	439	9	US-09-910-186A-14	Sequence 14, Appl
411	6	3.4	359	9	US-10-206-914-42	Sequence 42, Appl	484	6	3.4	456	10	US-09-997-664-92	Sequence 92, Appl
412	6	3.4	359	9	US-10-206-920-42	Sequence 42, Appl	485	6	3.4	458	9	US-10-174-590-496	Sequence 496, App
413	6	3.4	359	9	US-10-206-921-42	Sequence 42, Appl	486	6	3.4	458	9	US-10-176-758-496	Sequence 496, App
414	6	3.4	359	9	US-10-206-923-42	Sequence 42, Appl	487	6	3.4	458	9	US-10-175-737-496	Sequence 496, App
415	6	3.4	359	9	US-10-206-925-42	Sequence 42, Appl	488	6	3.4	458	9	US-10-173-706-496	Sequence 496, App
416	6	3.4	359	9	US-10-206-926-42	Sequence 42, Appl	489	6	3.4	458	9	US-10-175-738-496	Sequence 496, App
417	6	3.4	359	9	US-10-206-927-42	Sequence 42, Appl	490	6	3.4	458	9	US-10-175-752-496	Sequence 496, App
418	6	3.4	359	9	US-10-207-916-42	Sequence 42, Appl	491	6	3.4	458	9	US-10-176-482-496	Sequence 496, App
419	6	3.4	359	9	US-10-207-917-42	Sequence 42, Appl	492	6	3.4	458	9	US-10-176-757-496	Sequence 496, App
420	6	3.4	359	9	US-10-207-918-42	Sequence 42, Appl	493	6	3.4	458	9	US-10-176-913-496	Sequence 496, App
421	6	3.4	359	9	US-10-207-919-42	Sequence 42, Appl	494	6	3.4	458	9	US-10-180-552-496	Sequence 496, App
422	6	3.4	359	9	US-10-207-920-42	Sequence 42, Appl	495	6	3.4	458	9	US-10-180-557-496	Sequence 496, App
423	6	3.4	359	9	US-10-207-925-42	Sequence 42, Appl	496	6	3.4	458	9	US-10-173-700-496	Sequence 496, App
424	6	3.4	359	9	US-10-208-021-42	Sequence 42, Appl	497	6	3.4	458	9	US-10-174-572-496	Sequence 496, App
425	6	3.4	359	9	US-10-208-022-42	Sequence 42, Appl	498	6	3.4	458	9	US-10-174-579-496	Sequence 496, App
426	6	3.4	359	9	US-10-208-023-42	Sequence 42, Appl	499	6	3.4	458	9	US-10-174-582-496	Sequence 496, App
427	6	3.4	359	9	US-10-208-026-42	Sequence 42, Appl	500	6	3.4	458	9	US-10-174-588-496	Sequence 496, App
428	6	3.4	359	9	US-10-208-029-42	Sequence 42, Appl	501	6	3.4	458	9	US-10-175-739-496	Sequence 496, App
429	6	3.4	359	9	US-10-208-030-42	Sequence 42, Appl	502	6	3.4	458	9	US-10-175-740-496	Sequence 496, App
430	6	3.4	359	9	US-10-232-232-42	Sequence 42, Appl	503	6	3.4	458	9	US-10-175-743-496	Sequence 496, App
431	6	3.4	359	9	US-09-978-193A-45	Sequence 45, Appl	504	6	3.4	458	9	US-10-176-488-496	Sequence 496, App
432	6	3.4	359	9	US-10-013-929A-45	Sequence 45, Appl	505	6	3.4	458	9	US-10-176-492-496	Sequence 496, App
433	6	3.4	359	9	US-10-016-177A-45	Sequence 45, Appl	506	6	3.4	458	9	US-10-176-747-496	Sequence 496, App
434	6	3.4	359	9	US-10-173-693-42	Sequence 42, Appl	507	6	3.4	458	9	US-10-176-750-496	Sequence 496, App
435	6	3.4	359	9	US-10-174-578-42	Sequence 42, Appl	508	6	3.4	458	9	US-10-176-985-496	Sequence 496, App
436	6	3.4	359	9	US-10-175-741-42	Sequence 42, Appl	509	6	3.4	458	9	US-10-176-987-496	Sequence 496, App
437	6	3.4	359	9	US-10-175-750-42	Sequence 42, Appl	510	6	3.4	458	9	US-10-176-991-496	Sequence 496, App
438	6	3.4	359	9	US-10-176-986-42	Sequence 42, Appl	511	6	3.4	458	9	US-10-176-992-496	Sequence 496, App
439	6	3.4	359	9	US-10-184-641-42	Sequence 42, Appl	512	6	3.4	458	9	US-10-176-993-496	Sequence 496, App
440	6	3.4	359	9	US-10-187-888-42	Sequence 42, Appl	513	6	3.4	458	9	US-10-184-658-496	Sequence 496, App
441	6	3.4	359	9	US-10-194-360-42	Sequence 42, Appl	514	6	3.4	458	9	US-10-173-695-496	Sequence 496, App
442	6	3.4	359	9	US-10-194-365-42	Sequence 42, Appl	515	6	3.4	458	9	US-10-173-697-496	Sequence 496, App
443	6	3.4	359	9	US-10-195-895-42	Sequence 42, Appl	516	6	3.4	458	9	US-10-173-705-496	Sequence 496, App
444	6	3.4	359	9	US-10-195-898-42	Sequence 42, Appl	517	6	3.4	458	9	US-10-174-576-496	Sequence 496, App
445	6	3.4	359	9	US-10-196-759-42	Sequence 42, Appl	518	6	3.4	458	9	US-10-174-585-496	Sequence 496, App
446	6	3.4	359	9	US-10-199-302-42	Sequence 42, Appl	519	6	3.4	458	9	US-10-174-586-496	Sequence 496, App
447	6	3.4	359	9	US-10-201-323-42	Sequence 42, Appl	520	6	3.4	458	9	US-10-175-747-496	Sequence 496, App
448	6	3.4	359	9	US-10-205-510-42	Sequence 42, Appl	521	6	3.4	458	9	US-10-176-481-496	Sequence 496, App
449	6	3.4	359	9	US-10-205-891-42	Sequence 42, Appl	522	6	3.4	458	9	US-10-176-485-496	Sequence 496, App
450	6	3.4	359	9	US-10-205-904-42	Sequence 42, Appl	523	6	3.4	458	9	US-10-176-487-496	Sequence 496, App
451	6	3.4	359	9	US-10-206-917-42	Sequence 42, Appl	524	6	3.4	458	9	US-10-176-493-496	Sequence 496, App
452	6	3.4	359	9	US-10-207-923-42	Sequence 42, Appl	525	6	3.4	458	9	US-10-176-756-496	Sequence 496, App
453	6	3.4	359	9	US-10-207-924-42	Sequence 42, Appl	526	6	3.4	458	9	US-10-176-911-496	Sequence 496, App
454	6	3.4	359	9	US-10-208-028-42	Sequence 42, Appl	527	6	3.4	458	9	US-10-176-919-496	Sequence 496, App
455	6	3.4	359	9	US-09-999-830A-45	Sequence 45, Appl	528	6	3.4	458	9	US-10-176-925-496	Sequence 496, App
456	6	3.4	359	9	US-10-121-062-42	Sequence 42, Appl	529	6	3.4	458	9	US-10-176-978-496	Sequence 496, App
457	6	3.4	359	9	US-10-175-753-42	Sequence 42, Appl	530	6	3.4	458	9	US-10-175-510-496	Sequence 496, App

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:48:56 ; Search time 20 seconds
(without alignments)
823.630 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 179

Sequence: 1 QHNGVGPICKGSLTKLP.....ADYTTKADANLYGLENLYRE 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 349150 seqs, 92025710 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.9	61	9	US-09-996-140-111
2	7	3.9	63	9	US-09-996-140-134
3	7	3.9	93	9	US-09-738-626-6630
4	7	3.9	218	9	US-10-278-173-150
5	7	3.9	334	9	US-10-278-173-82
6	7	3.9	511	10	US-09-881-752A-110
7	7	3.9	795	9	US-09-738-626-3810
8	6	3.4	8	9	US-10-024-918-10
9	6	3.4	21	9	US-10-150-111-181
10	6	3.4	24	9	US-10-097-065-537
11	6	3.4	34	10	US-09-864-761-36829
12	6	3.4	37	10	US-09-864-761-40107
13	6	3.4	59	9	US-10-122-25
14	6	3.4	60	10	US-09-867-550-1856
15	6	3.4	64	10	US-09-864-761-34896
16	6	3.4	64	10	US-09-867-550-904
17	6	3.4	68	10	US-09-864-761-41976
18	6	3.4	68	10	US-09-864-761-46121
19	6	3.4	69	10	US-09-925-300-1548

20	6	3.4	76	9	US-10-079-854-110	Sequence 110, App
21	6	3.4	76	10	US-09-764-878-110	Sequence 110, App
22	6	3.4	80	10	US-09-893-737-12	Sequence 12, Appl
23	6	3.4	82	9	US-09-738-626-6220	Sequence 6220, Ap
24	6	3.4	90	9	US-09-866-050A-134	Sequence 134, App
25	6	3.4	90	10	US-09-864-761-43578	Sequence 43578, A
26	6	3.4	91	10	US-09-864-761-43033	Sequence 43033, A
27	6	3.4	93	10	US-09-867-550-612	Sequence 612, App
28	6	3.4	102	9	US-10-300-616-49	Sequence 49, Appl
29	6	3.4	104	10	US-09-864-761-41522	Sequence 41522, A
30	6	3.4	110	9	US-10-092-154-844	Sequence 844, App
31	6	3.4	110	10	US-09-764-847-844	Sequence 844, App
32	6	3.4	111	9	US-10-125-258-8	Sequence 8, Appl
33	6	3.4	132	9	US-09-738-626-4318	Sequence 4318, Ap
34	6	3.4	137	9	US-09-862-540-51	Sequence 51, Appl
35	6	3.4	142	9	US-10-102-806-799	Sequence 799, Appl
36	6	3.4	148	9	US-10-125-258-10	Sequence 10, Appl
37	6	3.4	149	9	US-09-983-802-196	Sequence 196, App
38	6	3.4	153	10	US-09-925-300-1688	Sequence 1688, Ap
39	6	3.4	162	1	US-08-781-986A-5220	Sequence 5220, Ap
40	6	3.4	168	9	US-09-981-876-184	Sequence 184, App
41	6	3.4	168	9	US-09-148-545-184	Sequence 184, App
42	6	3.4	175	9	US-10-150-111-173	Sequence 70, Appl
43	6	3.4	183	9	US-09-895-298-70	Sequence 831, App
44	6	3.4	191	9	US-10-092-154-831	Sequence 831, App
45	6	3.4	191	10	US-09-764-847-831	Sequence 831, App
46	6	3.4	196	9	US-10-125-258-4	Sequence 4, Appl
47	6	3.4	217	10	US-09-764-853-664	Sequence 664, App
48	6	3.4	223	9	US-10-097-063-325	Sequence 325, App
49	6	3.4	231	9	US-10-122-822-1	Sequence 1, Appl
50	6	3.4	233	9	US-10-125-258-18	Sequence 18, Appl
51	6	3.4	253	9	US-10-150-111-84	Sequence 84, Appl
52	6	3.4	262	9	US-10-235-674-14	Sequence 14, Appl
53	6	3.4	262	10	US-09-263-689-14	Sequence 14, Appl
54	6	3.4	276	10	US-09-763-272-134	Sequence 134, App
55	6	3.4	278	10	US-09-925-300-1323	Sequence 1323, Ap
56	6	3.4	282	9	US-10-150-111-172	Sequence 172, App
57	6	3.4	305	9	US-10-102-806-624	Sequence 624, App
58	6	3.4	325	9	US-09-738-626-6035	Sequence 6035, Ap
59	6	3.4	325	9	US-09-746-660A-4	Sequence 4, Appl
60	6	3.4	326	9	US-09-957-187-75	Sequence 75, Appl
61	6	3.4	329	9	US-10-159-901-33	Sequence 33, Appl
62	6	3.4	329	9	US-10-159-901-34	Sequence 34, Appl
63	6	3.4	329	9	US-10-159-901-55	Sequence 55, Appl
64	6	3.4	340	10	US-09-250-883-17	Sequence 17, Appl
65	6	3.4	359	9	US-09-978-295A-45	Sequence 45, Appl
66	6	3.4	359	9	US-09-978-697-45	Sequence 45, Appl
67	6	3.4	359	9	US-09-978-192A-45	Sequence 45, Appl
68	6	3.4	359	9	US-09-999-832A-45	Sequence 45, Appl
69	6	3.4	359	9	US-09-978-189-45	Sequence 45, Appl
70	6	3.4	359	9	US-10-174-590-42	Sequence 42, Appl
71	6	3.4	359	9	US-10-176-758-42	Sequence 42, Appl
72	6	3.4	359	9	US-10-175-737-42	Sequence 42, Appl
73	6	3.4	359	9	US-10-173-706-42	Sequence 42, Appl
74	6	3.4	359	9	US-10-175-738-42	Sequence 42, Appl
75	6	3.4	359	9	US-10-175-752-42	Sequence 42, Appl
76	6	3.4	359	9	US-10-176-482-42	Sequence 42, Appl
77	6	3.4	359	9	US-10-176-757-42	Sequence 42, Appl
78	6	3.4	359	9	US-10-176-913-42	Sequence 42, Appl
79	6	3.4	359	9	US-10-180-552-42	Sequence 42, Appl
80	6	3.4	359	9	US-10-180-557-42	Sequence 42, Appl
81	6	3.4	359	9	US-10-173-700-42	Sequence 42, Appl
82	6	3.4	359	9	US-10-174-572-42	Sequence 42, Appl
83	6	3.4	359	9	US-10-174-579-42	Sequence 42, Appl
84	6	3.4	359	9	US-10-174-582-42	Sequence 42, Appl
85	6	3.4	359	9	US-10-174-588-42	Sequence 42, Appl
86	6	3.4	359	9	US-10-175-739-42	Sequence 42, Appl
87	6	3.4	359	9	US-10-175-740-42	Sequence 42, Appl
88	6	3.4	359	9	US-10-175-743-42	Sequence 42, Appl
89	6	3.4	359	9	US-10-176-488-42	Sequence 42, Appl
90	6	3.4	359	9	US-10-176-492-42	Sequence 42, Appl
91	6	3.4	359	9	US-10-176-747-42	Sequence 42, Appl
92	6	3.4	359	9	US-10-176-750-42	Sequence 42, Appl

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:46:15 : Search time 17 Seconds
(without alignments)
1012.239 Million cell updates/sec

Title: us-09-995-493-52

Perfect score: 179

Sequence: 1 QNGVLGPYIGKSLTKLP.....ADYTKATANLYGLNLYRF 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	7.3	450	2 B40183	outer membrane pro
2	13	7.3	455	2 A40183	outer membrane pro
3	13	7.3	459	2 A28787	outer membrane pro
4	13	7.3	459	2 F64065	outer membrane pro
5	9	5.0	412	2 G82406	long-chain fatty a
6	8	4.5	72	2 H69448	heme d1 biosynthes
7	8	4.5	417	2 E90844	probable factor li
8	8	4.5	417	2 D85702	hypothetical prote
9	8	4.5	417	2 A64869	ychp protein - Esc
10	8	4.5	633	2 F75107	abc transporter, A
11	7	3.9	66	2 D98068	degenerate transpo
12	7	3.9	146	1 MYUY	myoglobin [validat
13	7	3.9	168	2 S00942	hypothetical prote
14	7	3.9	180	2 AB2849	hypothetical prote
15	7	3.9	218	2 F84748	probable AP2 domai
16	7	3.9	233	2 T02590	DNA binding protei
17	7	3.9	260	1 CRMS2	carbonate dehydrat
18	7	3.9	261	2 E86354	hypothetical prote
19	7	3.9	274	2 B71877	hypothetical prote
20	7	3.9	308	2 T35643	selenium donor pro
21	7	3.9	336	2 F64498	glycerol-3-phospha
22	7	3.9	356	2 B30338	late competence pr
23	7	3.9	363	2 C84179	aminopeptidase hom
24	7	3.9	413	2 A70795	hypothetical prote
25	7	3.9	423	2 AH0334	probable long-chain
26	7	3.9	437	2 AE0805	long-chain fatty a
27	7	3.9	448	2 F65007	long-chain fatty a
28	7	3.9	448	2 C91032	long-chain fatty a
29	7	3.9	448	2 D85876	long-chain fatty a

probable outer mem
hypothetical prote
oligopeptide trans
hypothetical prote
iron(III)-transpor
B-lymphocyte antig
fiber protein - hu
B-cell antigen CD1
DNA ligase homolog
hypothetical prote
hypothetical prote
probable membrane
DNA-directed RNA p
outer membrane ush
alpha-adaptin A -
KIAA1688 protein l
conserved hypothet
kinase-related pro
hypothetical prote
hypothetical prote
Ig mu chain transm
zinc finger protei
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
slxX protein VC035
30S ribosomal prot
hypothetical prote
hypothetical prote
hypothetical prote
ribosomal protein
probable ribosomal
protein F21F23.5 l
transposase IS240-
hypothetical prote
ribosomal protein
fimbriata-associat
hypothetical prote
sporozoite antigen
phage-related prot
transcription regu
type E neurotoxin
ribosomal protein
ribosomal protein
ribosomal protein
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
probable kinetecho
SAPI protein homol
hypothetical prote
conserved hypothet
probable heat shoc
hypothetical prote
hypothetical prote
hypothetical prote
HD family hydrolas
lutropin receptor
hypothetical prote
outer surface prot
outer surface prot
conserved hypothet
CT469 hypothetical
hypothetical prote

103	6	3.4	184	2	ABI731	hypothetical prote	176	6	3.4	259	2	B85534	rRNA methylase [lm
104	6	3.4	185	2	D83658	stage v sporulatio	177	6	3.4	259	2	G85363	hypothetical prote
105	6	3.4	186	2	S14840	dihydrofolate redu	178	6	3.4	260	2	A95927	probable transcrip
106	6	3.4	187	1	AGEC2	p-aminobenzoate sy	179	6	3.4	261	2	T35708	hydrolase - Strept
107	6	3.4	187	2	S09636	pABA protein - Sai	180	6	3.4	261	2	T72426	iron-sulfur cluste
108	6	3.4	187	2	C91155	p-aminobenzoate sy	181	6	3.4	262	2	D71068	probable pseudouri
109	6	3.4	187	2	H86000	p-aminobenzoate sy	182	6	3.4	262	2	A54889	transcription regu
110	6	3.4	187	2	AB1002	p-aminobenzoate sy	183	6	3.4	263	2	C64944	probable regulator
111	6	3.4	188	2	D75499	probable acetyltra	184	6	3.4	263	2	A98946	icir-family transc
112	6	3.4	188	2	S70269	outer surface prot	185	6	3.4	263	2	AB0209	probable regulator
113	6	3.4	189	2	E90408	conserved hypoteth	186	6	3.4	263	2	E85794	transcription regu
114	6	3.4	189	2	B72348	hypothetical prote	187	6	3.4	263	2	AB0728	hypothetical prote
115	6	3.4	190	2	H75478	conserved hypoteth	188	6	3.4	263	2	T32081	hypothetical prote
116	6	3.4	190	2	G69339	conserved hypoteth	189	6	3.4	264	2	A28651	lactose-specific
117	6	3.4	191	2	G71710	sco2 protein precu	190	6	3.4	264	2	A45983	transcription regu
118	6	3.4	191	2	S70271	outer surface prot	191	6	3.4	264	2	AG3466	hypothetical prote
119	6	3.4	191	2	S70278	outer surface prot	192	6	3.4	266	2	C70788	transcription regu
120	6	3.4	191	2	H69387	conserved hypoteth	193	6	3.4	266	2	T37689	hypothetical prote
121	6	3.4	192	2	AD1036	probable phage bas	194	6	3.4	267	1	S33964	orotidine-5'-phos
122	6	3.4	192	2	AF0928	probable phage bas	195	6	3.4	267	2	D95089	hypothetical prote
123	6	3.4	194	2	T06831	endopeptidase Clp	196	6	3.4	267	2	G97956	hypothetical prote
124	6	3.4	194	2	S70277	outer surface prot	197	6	3.4	269	2	F95106	peptidyl-prolyl ci
125	6	3.4	195	2	T35651	probable methylase	198	6	3.4	269	2	H97974	Cof family protein
126	6	3.4	196	2	G83757	ribonuclease H-rel	199	6	3.4	270	2	G84226	conserved hypoteth
127	6	3.4	199	2	B97705	sco2 protein precu	200	6	3.4	270	2	T37809	hypothetical prote
128	6	3.4	203	2	S54224	outer surface prot	201	6	3.4	272	2	S33622	probable zinc flng
129	6	3.4	204	2	B87704	heme exporter prot	202	6	3.4	272	2	C84273	ADR6 protein - soy
130	6	3.4	204	2	B23700	hypothetical prote	203	6	3.4	277	2	A99517	ABC transporter at
131	6	3.4	206	2	E81221	probable integral	204	6	3.4	279	2	F97024	hypothetical prote
132	6	3.4	206	2	S76502	hypothetical prote	205	6	3.4	281	2	T35932	probable inositol
133	6	3.4	207	2	S71452	trypsin/chymotryps	206	6	3.4	281	2	T04858	hypothetical prote
134	6	3.4	207	2	S69924	outer surface prot	207	6	3.4	283	2	H64764	probable taurine d
135	6	3.4	208	2	C90551	hypothetical prote	208	6	3.4	283	2	F90681	taurine dioxygenas
136	6	3.4	208	2	H91033	probable resolvase	209	6	3.4	283	2	B85532	high-affinity tran
137	6	3.4	208	2	A85878	probable resolvase	210	6	3.4	283	2	F98243	ABC transporter, m
138	6	3.4	209	2	I40273	outer surface prot	211	6	3.4	283	2	AF3042	hypothetical prote
139	6	3.4	209	2	I40270	outer surface prot	212	6	3.4	285	2	G72414	hypothetical prote
140	6	3.4	209	2	T34027	hypothetical prote	213	6	3.4	288	1	B48583	spore germination
141	6	3.4	211	2	I40277	outer surface prot	214	6	3.4	288	2	T32128	hypothetical prote
142	6	3.4	211	2	S69928	outer surface prot	215	6	3.4	289	2	S35022	YabG protein - Bac
143	6	3.4	212	2	T06015	hypothetical prote	216	6	3.4	290	2	S66072	probable 3',5'-cyc
144	6	3.4	213	1	RD8EHS	dihydrofolate redu	217	6	3.4	290	2	G36789	probable 3',5'-cyc
145	6	3.4	216	1	S74616	uracil phosphoribo	218	6	3.4	291	1	F36516	uracil phosphoribo
146	6	3.4	216	2	S05575	sporozoite antigen	219	6	3.4	291	2	A12468	3-methylcatechol 2
147	6	3.4	217	2	S61625	hypothetical prote	220	6	3.4	294	2	D90578	hypothetical prote
148	6	3.4	224	2	T03969	uracil phosphoribo	221	6	3.4	295	1	D72021	elongation factor
149	6	3.4	228	2	D64672	ABC transporter, A	222	6	3.4	295	1	G86602	probable 3',5'-cyc
150	6	3.4	228	2	G71842	probable ABC trans	223	6	3.4	296	2	T45930	probable 3',5'-cyc
151	6	3.4	228	2	G64573	protein phosphatas	224	6	3.4	296	2	B49889	nonstructural prot
152	6	3.4	229	2	A82959	hypothetical prote	225	6	3.4	297	1	MNNAV	salivary glue prot
153	6	3.4	233	1	EWKBC	attacin precursor,	226	6	3.4	297	2	S1085	nonstructural phos
154	6	3.4	234	2	T07686	transcription fact	227	6	3.4	299	2	A95878	probable N-acetyl
155	6	3.4	234	2	A75587	probable urea/short	228	6	3.4	299	2	AG2340	hypothetical prote
156	6	3.4	236	2	T02432	ethylene-responsiv	229	6	3.4	301	2	D82040	cyso protein VC272
157	6	3.4	236	2	AF2288	hypothetical prote	230	6	3.4	302	2	B49889	regulatory protein
158	6	3.4	239	2	T36870	probable DNA hydro	231	6	3.4	303	2	T04541	hypothetical prote
159	6	3.4	242	2	JC4300	galactin-3 - rabbi	232	6	3.4	304	2	T34845	hypothetical prote
160	6	3.4	243	2	S06098	hypothetical 26.6K	233	6	3.4	305	1	H75285	probable phosphos
161	6	3.4	243	2	H82253	ribonuclease T VC1	234	6	3.4	305	2	T03155	ribonucleoside-dip
162	6	3.4	244	2	AB3524	transcription regu	235	6	3.4	305	2	T08836	probable metallop
163	6	3.4	245	2	A54909	carbohydrate-bind	236	6	3.4	305	2	A10476	oxidative stress t
164	6	3.4	245	2	G75422	hypothetical prote	237	6	3.4	306	2	S17713	transcription regu
165	6	3.4	247	1	S18532	hypothetical prote	238	6	3.4	308	2	A84566	Ac-like transposas
166	6	3.4	247	1	G84280	probable phosphos	239	6	3.4	308	2	T46023	hypothetical prote
167	6	3.4	247	1	A55910	H-transporing tw	240	6	3.4	309	2	T21195	hypothetical prote
168	6	3.4	250	2	B71859	probable oxidoredu	241	6	3.4	310	2	T36542	hypothetical prote
169	6	3.4	250	2	E64564	short chain alcoh	242	6	3.4	313	1	H69510	conserved hypoteth
170	6	3.4	253	2	AB3034	probable short-cha	243	6	3.4	313	2	JC5475	pollen allergen 9
171	6	3.4	253	2	H70217	conserved hypoteth	244	6	3.4	314	2	AF3268	signal recognition
172	6	3.4	254	2	AG3331	cobalt transport p	245	6	3.4	314	2	T23301	hypothetical prote
173	6	3.4	257	2	T23332	hypothetical prote	246	6	3.4	317	2	JC5696	prolyl aminopeptid
174	6	3.4	258	2	S75950	hypothetical prote	247	6	3.4	319	2	A70959	probable aspartate
175	6	3.4	259	2	B72089	ABC transporter, A	248	6	3.4	320	2	T36062	probable ABC-type

249	6	3.4	321	2	D86975	probable aspartate	322	6	3.4	401	2	E84367	alcohol dehydrogen
250	6	3.4	321	2	S54403	1-phosphatidylinos	323	6	3.4	404	2	H90983	probable glycosyl
251	6	3.4	323	2	T05478	peroxidase (EC 1.1	324	6	3.4	404	2	C85829	glycosyl transfera
252	6	3.4	324	2	F69504	proliferating-cell	325	6	3.4	405	2	A55967	1-phosphatidylinos
253	6	3.4	325	2	T05356	hypothetical prote	326	6	3.4	406	2	S72177	1-phosphatidylinos
254	6	3.4	326	2	A41862	C-S lyase (ORF326)	327	6	3.4	406	2	E81300	probable glucose-6
255	6	3.4	327	2	A64613	conserved hypother	328	6	3.4	407	2	B71690	hypothetical prote
256	6	3.4	327	2	B71900	hypothetical prote	329	6	3.4	411	2	B71500	probable amino aci
257	6	3.4	329	1	JC4251	D-xylose 1-dehydro	330	6	3.4	413	2	A97542	D-ala-D-ala-carbox
258	6	3.4	331	2	A99179	conserved hypother	331	6	3.4	414	2	D86484	hypothetical prote
259	6	3.4	333	2	G87426	ADP-L-glycero-D-ma	332	6	3.4	415	2	B86434	protein T17H7.13 l
260	6	3.4	333	2	T40923	pyridoxal reductas	333	6	3.4	418	2	H83128	probable porin PA4
261	6	3.4	334	2	T19127	hypothetical prote	334	6	3.4	420	2	C69128	glutamate-1-semial
262	6	3.4	335	2	D84332	hypB protein - Rho	335	6	3.4	420	2	I49708	GTP-binding protei
263	6	3.4	336	2	T49112	aspartic proteinas	336	6	3.4	422	2	A12563	hypothetical prote
264	6	3.4	337	2	C84297	hypothetical prote	337	6	3.4	423	2	I59463	keratin, type I, c
265	6	3.4	337	2	AH3207	transcription regu	338	6	3.4	424	2	B81936	tetrahydrofolylpol
266	6	3.4	345	2	B97066	aldose-1-epimerase	339	6	3.4	424	2	C81170	folypolyglutamate
267	6	3.4	345	2	T48313	hypothetical prote	340	6	3.4	424	2	S63583	tetrahydrofolylpol
268	6	3.4	347	2	B85432	hypothetical prote	341	6	3.4	424	2	T43498	hypothetical prote
269	6	3.4	349	2	D70141	oligopeptide ABC t	342	6	3.4	425	2	AG2787	long-chain fatty a
270	6	3.4	350	2	C96532	peroxidase Atp5a l	343	6	3.4	428	2	E87573	chlorohydrolase l
271	6	3.4	352	2	A49210	heat shock:protein	344	6	3.4	429	2	A97567	hypothetical prote
272	6	3.4	352	2	C82294	oxidoreductase fas	345	6	3.4	430	2	T34627	probable electron
273	6	3.4	352	2	T18794	hypothetical prote	346	6	3.4	430	2	E83870	hypothetical prote
274	6	3.4	356	2	F82069	conserved hypother	347	6	3.4	431	2	T24101	hypothetical prote
275	6	3.4	356	2	D82643	conserved hypother	348	6	3.4	431	2	E82766	conserved hypother
276	6	3.4	357	2	F82878	XAA-PRO aminopepti	349	6	3.4	432	2	B82250	long-chain fatty a
277	6	3.4	359	2	A75438	recF protein - Del	350	6	3.4	432	2	A85436	APETALA2 protein l
278	6	3.4	360	2	I51663	arginase 1 - Afric	351	6	3.4	432	2	E87508	outer membrane pro
279	6	3.4	360	2	I51664	arginase 2 - Afric	352	6	3.4	433	2	D82879	ABC transporter DU
280	6	3.4	360	2	I51665	arginase 3 - Afric	353	6	3.4	433	2	D83456	probable guanine d
281	6	3.4	362	2	C64807	ybgO protein - Esc	354	6	3.4	434	2	T28342	ORF MSV181 hypoth
282	6	3.4	362	2	T29552	hypothetical prote	355	6	3.4	442	2	D95889	conserved hypother
283	6	3.4	363	2	G11544	hypothetical prote	356	6	3.4	444	2	PD0001	protein-glutamine
284	6	3.4	364	2	D70164	heat shock protein	357	6	3.4	445	2	T24105	hypothetical prote
285	6	3.4	365	2	E83710	hypothetical prote	358	6	3.4	446	2	T03981	APETALA2-like prot
286	6	3.4	366	2	T48501	hypothetical prote	359	6	3.4	447	2	T02547	hypothetical prote
287	6	3.4	368	1	I64231	high affinity tran	360	6	3.4	447	2	C75413	probable proteinas
288	6	3.4	369	2	E75620	hypothetical prote	361	6	3.4	448	2	F84780	hypothetical prote
289	6	3.4	370	2	E82369	thiH protein VC006	362	6	3.4	448	2	F95122	protein kinase, pr
290	6	3.4	371	2	A71378	probable DNA polym	363	6	3.4	448	2	G70977	hypothetical prote
291	6	3.4	371	2	F70634	hypothetical prote	364	6	3.4	450	2	E70681	probable trpE - My
292	6	3.4	371	2	AH2172	hypothetical prote	365	6	3.4	453	2	G96573	protein F12M16.27
293	6	3.4	373	2	D64106	dprA protein - Hae	366	6	3.4	453	2	S53995	TodX protein - Pse
294	6	3.4	374	2	JC2124	major allergen Cry	367	6	3.4	454	2	T24455	hypothetical prote
295	6	3.4	374	2	JC2123	major allergen Cry	368	6	3.4	454	2	G97125	probable inactivat
296	6	3.4	377	2	T47439	hypothetical prote	369	6	3.4	455	2	S71344	purinergic recepto
297	6	3.4	378	2	E70786	hypothetical prote	370	6	3.4	460	2	B34283	H+-transporting tw
298	6	3.4	383	2	D86781	L-lactate oxidase	371	6	3.4	460	2	T45108	H+-transporting tw
299	6	3.4	383	2	S56906	GCPI4 protein - ye	372	6	3.4	462	2	C83747	dihydroliipoamide d
300	6	3.4	386	2	AD0098	probable sugar tra	373	6	3.4	464	2	C84686	probable AP2 domai
301	6	3.4	386	2	B97035	uncharacterized pr	374	6	3.4	467	2	T34874	hypothetical prote
302	6	3.4	387	2	AH3624	spermidine/putresc	375	6	3.4	468	2	JC4285	acid phosphatase (
303	6	3.4	387	2	A19498	DNAJ protein - Str	376	6	3.4	469	2	JC5798	F0F1-ATPase (EC 3.
304	6	3.4	388	2	B75557	probable coproporp	377	6	3.4	469	2	T48314	hypothetical prote
305	6	3.4	388	2	C83102	probable MFS trans	378	6	3.4	471	2	E83130	conserved hypother
306	6	3.4	390	2	E69272	conserved hypother	379	6	3.4	475	2	D90605	hexosephosphate tr
307	6	3.4	390	2	AC2761	penicillin binding	380	6	3.4	475	2	T18487	hypothetical prote
308	6	3.4	391	2	AG0533	probable drug effi	381	6	3.4	476	2	C72016	glycogen synthase
309	6	3.4	391	2	S21318	hypothetical prote	382	6	3.4	476	2	B86609	glycogen synthase
310	6	3.4	391	2	AH1762	hypothetical prote	383	6	3.4	478	2	D97682	hypothetical prote
311	6	3.4	391	2	G90253	formate hydrogenly	384	6	3.4	478	2	AD2907	cell division part
312	6	3.4	393	2	S74688	hypothetical prote	385	6	3.4	480	2	T18905	hypothetical prote
313	6	3.4	393	2	A64036	hypothetical prote	386	6	3.4	480	2	T18675	hypothetical prote
314	6	3.4	394	2	C85064	hypothetical prote	387	6	3.4	481	2	JC5378	protein disulfide-
315	6	3.4	394	2	S50696	hypothetical prote	388	6	3.4	483	2	T23312	hypothetical prote
316	6	3.4	396	1	C69988	chloramphenicol re	389	6	3.4	484	2	F83005	thiazole biosynthe
317	6	3.4	397	2	B75592	UDP-galactopyranos	390	6	3.4	486	2	S43274	gag-like protein -
318	6	3.4	399	2	S70206	dnaJ protein - Str	391	6	3.4	487	2	S75338	NADH2 dehydrogenas
319	6	3.4	399	2	H87486	1-deoxy-D-xyulose	392	6	3.4	488	2	I51379	synthrophin - Pacif
320	6	3.4	400	2	AH0666	probable glycolate	393	6	3.4	490	2	T09084	phosphatidylinosit
321	6	3.4	401	2	H70658	probable aroF prot	394	6	3.4	491	2	E95082	lysyl-tRNA synthet

395	6	3.4	491	2	S73110	hypothetical prote	468	6	3.4	615	2	G90575	transketolase (Lk)
396	6	3.4	493	2	S11148	amiA protein - Str	469	6	3.4	617	2	B46091	H ⁺ -exporting ATPase
397	6	3.4	494	2	E86671	lysine-tRNA ligase	470	6	3.4	619	2	T44285	heat-shock-cognate
398	6	3.4	494	2	D86493	polymorphic outer	471	6	3.4	620	2	A70525	hypothetical prote
399	6	3.4	496	2	B97950	lysine-tRNA ligase	472	6	3.4	621	2	D71961	glucose inhibited
400	6	3.4	498	2	F98149	hypothetical prote	473	6	3.4	621	2	E64546	glucose inhibited
401	6	3.4	499	2	AH0485	probable insulinase	474	6	3.4	621	2	A83421	peptidyl-prolyl ci
402	6	3.4	502	2	F90003	ATP synthase alpha	475	6	3.4	622	2	D44986	apical membrane an
403	6	3.4	503	2	AC3148	exopolysaccharide	476	6	3.4	622	2	B44986	apical membrane an
404	6	3.4	505	2	F90427	amino acid transpo	477	6	3.4	622	2	A32499	apical membrane an
405	6	3.4	506	2	C95217	choline transporte	478	6	3.4	622	2	A44986	apical membrane an
406	6	3.4	506	2	T44634	choline transporte	479	6	3.4	622	2	C44986	apical membrane an
407	6	3.4	506	2	C98081	hypothetical prote	480	6	3.4	627	2	G84558	probable SEC1 fami
408	6	3.4	509	2	T30861	traj protein - Sai	481	6	3.4	633	2	H84336	Htr3 transducer [1
409	6	3.4	510	2	S53958	hypothetical prote	482	6	3.4	633	2	G71026	hypothetical prote
410	6	3.4	512	2	T18803	hypothetical prote	483	6	3.4	634	2	F96652	protein F23N19.13
411	6	3.4	514	2	S18449	variant surface gl	484	6	3.4	634	2	T16897	hypothetical prote
412	6	3.4	514	2	S56384	hypothetical 56.3k	485	6	3.4	635	2	T46407	probable RNA helic
413	6	3.4	517	2	B71260	hypothetical prote	486	6	3.4	637	2	I49638	probable RNA helic
414	6	3.4	518	2	H86826	signal recognition	487	6	3.4	638	2	T28932	probable 4-coumara
415	6	3.4	520	2	AD2616	chromosomal replic	488	6	3.4	644	1	H69204	endopeptidase La h
416	6	3.4	520	2	S72324	orotate reductase	489	6	3.4	644	2	A86623	transketolase [imp
417	6	3.4	525	2	T00660	hypothetical prote	490	6	3.4	644	2	A72002	1-deoxyxylulose-5-
418	6	3.4	529	2	D97398	dnaA protein (L254	491	6	3.4	645	2	T11137	NADH2 dehydrogenas
419	6	3.4	529	2	A95423	probable ABC trans	492	6	3.4	645	2	G81315	membrane bound zin
420	6	3.4	532	2	G83424	hypothetical prote	493	6	3.4	653	2	F83405	probable TonB-depe
421	6	3.4	533	2	T04481	Mlo protein - bari	494	6	3.4	657	2	G86397	protein T7N9.18 [1
422	6	3.4	533	2	AE3138	hypothetical prote	495	6	3.4	658	1	A39018	carnitine O-palmit
423	6	3.4	533	2	T40153	hypothetical prote	496	6	3.4	658	2	A49362	carnitine O-palmit
424	6	3.4	535	2	A87570	ABC transporter, A	497	6	3.4	659	2	A95221	hypothetical prote
425	6	3.4	537	2	S39970	glucan 1,6-alpha-g	498	6	3.4	659	2	A98085	hypothetical prote
426	6	3.4	537	2	A38170	cytochrome d compl	499	6	3.4	662	2	A37226	glucose transport
427	6	3.4	540	2	H86777	hypothetical prote	500	6	3.4	663	2	C82415	ATP-dependent RNA
428	6	3.4	542	2	C86405	probable sphingosl	501	6	3.4	668	2	G83349	conserved hypoteth
429	6	3.4	547	2	S49815	transferrin-bindin	502	6	3.4	672	2	D81746	type III secretion
430	6	3.4	547	2	A47966	transferrin-bindin	503	6	3.4	676	2	A45900	complement C3b rec
431	6	3.4	548	2	A86385	hypothetical prote	504	6	3.4	676	2	AF1153	transcription anti
432	6	3.4	550	2	H98139	gumJ protein [impo	505	6	3.4	685	2	AE3408	exonuclease ABC c
433	6	3.4	551	2	JC7562	glioblastoma RING	506	6	3.4	687	1	B39405	protein-glutamine
434	6	3.4	551	2	T23945	hypothetical prote	507	6	3.4	690	2	A25534	hypothetical prote
435	6	3.4	556	1	A55483	transcription init	508	6	3.4	691	1	A29996	protein-glutamine
436	6	3.4	557	1	NUVKL	glucose-6-phosphat	509	6	3.4	691	2	C89935	penicillin-binding
437	6	3.4	557	2	D83478	gamma-glutamyltran	510	6	3.4	693	2	A45991	protein-glutamine
438	6	3.4	557	2	T45135	chaperone protein	511	6	3.4	693	2	B45991	protein-glutamine
439	6	3.4	558	2	C72391	conserved hypoteth	512	6	3.4	698	1	S31714	prqtein-glutamine
440	6	3.4	558	2	AE3580	iron(III)-transpor	513	6	3.4	698	1	S31630	NAD+-protein ADP-r
441	6	3.4	561	2	T01950	hypothetical prote	514	6	3.4	698	2	F83096	probable two-compo
442	6	3.4	562	2	AH2084	hypothetical prote	515	6	3.4	698	2	E82443	heme transport pro
443	6	3.4	563	2	S32156	mandelonitrile lya	516	6	3.4	702	2	H84189	lipocate protein li
444	6	3.4	568	1	D64059	2-succinyl-6-hydro	517	6	3.4	708	2	I83196	NEDD-4 ORF - mouse
445	6	3.4	570	2	H90370	thermopsin precurs	518	6	3.4	708	2	AH2626	ferrichrome iron r
446	6	3.4	571	2	T10232	hypothetical prote	519	6	3.4	713	2	JC2534	RVLG protein - rat
447	6	3.4	572	2	C75515	conserved hypoteth	520	6	3.4	721	2	C84073	DNA topoisomerase
448	6	3.4	572	2	T08509	trbl protein - Ent	521	6	3.4	722	2	F71270	probable flagellar
449	6	3.4	574	2	G69769	pyruvate oxidase h	522	6	3.4	724	2	T27858	probable peroxidase
450	6	3.4	575	1	S35138	acetylactate synth	523	6	3.4	724	2	S57604	probable membrane
451	6	3.4	577	2	D83473	isocitrate dehydro	524	6	3.4	728	2	F97262	cupulsomerase B [1
452	6	3.4	582	2	H81671	single-stranded-DN	525	6	3.4	731	1	A55800	cucumisin (BC 3.4.
453	6	3.4	582	2	C59099	hypothetical prote	526	6	3.4	738	2	T00343	hypothetical prote
454	6	3.4	582	2	A55174	kinase-associated	527	6	3.4	739	2	T10648	hypothetical prote
455	6	3.4	583	2	H84810	probable calcium-d	528	6	3.4	744	2	C75276	sensory box sensor
456	6	3.4	584	2	H71513	probable ssdna exo	529	6	3.4	747	2	G97408	fegA protein U6140
457	6	3.4	585	2	S48929	hypothetical prote	530	6	3.4	751	2	T01449	cytoskeletal prote
458	6	3.4	587	2	A31776	lactose permease -	531	6	3.4	759	2	T24130	hypothetical prote
459	6	3.4	595	2	T06111	hypothetical prote	532	6	3.4	759	2	B70585	hypothetical prote
460	6	3.4	596	1	A44072	transcription fact	533	6	3.4	763	2	A43747	X-Pro dipeptidyl-p
461	6	3.4	596	2	I40609	DNA primase (EC 2.	534	6	3.4	763	2	A43748	X-pro dipeptidyl-p
462	6	3.4	596	2	C97060	DNA primase, DNAG	535	6	3.4	763	2	A86881	hypothetical prote
463	6	3.4	598	2	I39596	chitinobiose - Alter	536	6	3.4	763	2	T22843	hypothetical prote
464	6	3.4	607	2	S48326	hypothetical prote	537	6	3.4	767	2	S63220	probable membrane
465	6	3.4	610	2	C96732	hypothetical prote	538	6	3.4	780	2	T30184	hypothetical prote
466	6	3.4	611	2	A85000	heat shock protein	539	6	3.4	787	2	T16901	hypothetical prote
467	6	3.4	614	1	S61532	RET oncogene fusio	540	6	3.4	787	2	S56268	hypothetical prote

541	6	3.4	791	2	F81056	hemoglobin recepto	614	1251	2	JH0256	botulinum neurotox
542	6	3.4	792	2	S61335	hemoglobin recepto	615	1252	2	S21178	botulinum neurotox
543	6	3.4	793	1	KXRTF	furin (EC 3.4.21.7	616	1286	1	RJBOP	interphotoreceptor
544	6	3.4	796	1	A32434	absa protein - Eme	617	1290	2	S73653	DNA-directed RNA p
545	6	3.4	803	2	AB0530	outer membrane pro	618	1297	2	S39791	neurotoxin - Clost
546	6	3.4	805	2	T48840	transducer protein	619	1301	2	D85188	disease resistance
547	6	3.4	806	2	E64221	phenylalanine-tRNA	620	1305	2	T12897	DNA polymerase III
548	6	3.4	806	2	T08932	hypothetical prote	621	1338	2	A49634	aldehyde oxidase (
549	6	3.4	809	2	T41645	probable spindle p	622	1341	2	H98323	hypothetical prote
550	6	3.4	811	2	B82365	glycerol-3-phospha	623	1351	2	E97273	superfamily I DNA
551	6	3.4	812	1	MECOF	outer membrane ush	624	1355	2	T00961	hypothetical prote
552	6	3.4	815	2	AB2444	hypothetical prote	625	1375	2	JT0345	dextranucrase (EC
553	6	3.4	819	2	A13197	TomB-dependent rec	626	1400	2	T31555	hypothetical prote
554	6	3.4	822	2	T33163	hypothetical prote	627	1430	2	AF0351	probable autotrans
555	6	3.4	826	2	D86458	hypothetical prote	628	1430	2	T13018	hypothetical prote
556	6	3.4	828	2	T03544	hypothetical prote	629	1446	2	B84426	hypothetical prote
557	6	3.4	833	1	S20387	outer membrane pro	630	1449	2	A97323	DNA polymerase III
558	6	3.4	833	2	T14703	F1 capsule anchori	631	1452	2	S73223	genome polyprotein
559	6	3.4	839	2	B96576	hypothetical prote	632	1470	2	A45323	C-terminal domain-
560	6	3.4	843	2	S78372	DNA-directed RNA p	633	1473	2	T31422	9tF protein precu
561	6	3.4	854	2	T23155	hypothetical prote	634	1475	2	B33135	cystic fibrosis tr
562	6	3.4	858	2	A44919	GCR3 protein - yea	635	1476	1	A39901	cystic fibrosis tr
563	6	3.4	859	2	G86242	hypothetical prote	636	1476	1	A40303	alpha-macroglobuli
564	6	3.4	860	2	T28227	ORF MSV067 probabl	637	1476	2	JC5143	multiresistance p
565	6	3.4	870	2	AD0625	aminopeptidase N I	638	1514	2	T52080	hypothetical prote
566	6	3.4	873	2	S75028	hypothetical prote	639	1532	2	T18438	sol protein, large
567	6	3.4	875	2	T17382	vric protein - Dic	640	1597	1	BVFESL	gene small optic l
568	6	3.4	875	2	T27495	hypothetical prote	641	1597	2	T08428	hypothetical prote
569	6	3.4	887	2	S70642	ubiquitin ligase N	642	1638	2	T25352	hypothetical prote
570	6	3.4	903	2	D87250	DNA mismatch repai	643	1652	2	T16799	probable membrane
571	6	3.4	905	2	T23229	hypothetical prote	644	1679	2	S49802	hypothetical prote
572	6	3.4	906	1	RNBPK1	DNA-directed RNA p	645	1755	2	T51532	probable polyketid
573	6	3.4	908	2	JN0819	transferrin-bindin	646	1762	2	T03222	hypothetical prote
574	6	3.4	916	2	H69161	DNA helicase II -	647	1817	2	T10689	probable polyketid
575	6	3.4	919	1	QVFKG	phosphoenolpyruvat	648	1937	2	T03224	enoyl-[acyl-carrie
576	6	3.4	936	2	C86546	polymorphic outer	649	2051	2	S34688	fatty acyl-CoA syn
577	6	3.4	936	2	B81591	polymorphic membra	650	2073	2	T43311	fatty acid synthas
578	6	3.4	940	2	C72078	polymorphic outer	651	2073	2	T39207	rough deal protein
579	6	3.4	946	2	H71409	hypothetical prote	652	2098	2	T13166	probable polyketid
580	6	3.4	946	2	F84280	ATP-dependent hell	653	2100	2	T03223	probable polyketid
581	6	3.4	953	1	B30169	leukotoxin A - Pas	654	2133	2	T42763	coagulation factor
582	6	3.4	954	2	A10438	probable exported	655	2149	2	C96695	ribulose bispoph
583	6	3.4	960	2	JE0336	gamma-aminobutyric	656	2214	2	T18305	hypothetical prote
584	6	3.4	982	2	T34830	polyketide synthas	657	2319	2	A47004	coagulation factor
585	6	3.4	988	2	T51054	related to alpha-a	658	2351	1	EZH0	coagulation factor
586	6	3.4	992	1	GNWVR3	structural polypro	659	3014	1	JC5620	genome polyprotein
587	6	3.4	992	2	J01165	Env protein - Maed	660	3078	2	T28432	variant-specific s
588	6	3.4	996	2	E98092	cylM protein, cyto	661	3268	2	S69625	hypothetical prote
589	6	3.4	996	2	A95228	bacteriocin format	662	3433	1	S28381	utrophin - human
590	6	3.4	1004	2	JC2221	major surface glyc	663	3512	2	T17121	Cpr protein - midg
591	6	3.4	1014	1	NMGLSS	exo-alpha-sialidas	664	3512	2	T17121	dynain heavy chain
592	6	3.4	1018	2	A54744	contactin 1 precu	665	4092	1	S38128	MEGF1 protein - ra
593	6	3.4	1032	2	T40658	scrine-threonine p	666	4436	2	E71086	hypothetical prote
594	6	3.4	1037	2	T27345	hypothetical prote	667	4589	2	T14914	dynain beta heavy
595	6	3.4	1038	2	S36801	pullulanase (EC 3.	668	4613	2	T17409	polyketide synthas
596	6	3.4	1049	2	C95883	probable efflux pr	669	4639	1	A54794	dynain heavy chain
597	6	3.4	1052	2	AF2959	conserved hypothet	670	4735	2	T17463	rifamycin polyketi
598	6	3.4	1054	2	T43226	translation initia	671	4881	2	T18489	hypothetical prote
599	6	3.4	1058	2	AG2541	cation efflux syst	672	6831	2	A88852	protein unc-22 [im
600	6	3.4	1063	1	GNWV77	structural polypro	673	6839	2	S57242	twlchlin [similari
601	6	3.4	1073	2	T19745	hypothetical prote	674	7160	2	T27935	hypothetical prote
602	6	3.4	1090	2	S11823	pullulanase (EC 3.	675	8563	2	T30226	polyketide synthas
603	6	3.4	1096	2	A26879	pullulanase (EC 3.	676	10223	2	T30225	polyketide synthas
604	6	3.4	1118	2	S48383	probable membrane	677	13	2	S36874	cytochrome P450 CM
605	6	3.4	1154	2	T19525	diacylglycerol kin	678	14	2	A47421	leukotriene B-4 12
606	6	3.4	1173	2	T31421	C-terminal domain-	679	18	2	A43826	dnak-type molecula
607	6	3.4	1184	2	H71436	hypothetical prote	680	20	2	S21737	coumarin 7-monooxy
608	6	3.4	1195	2	S61886	chitin synthase (E	681	20	2	S02002	photosystem II pro
609	6	3.4	1217	2	H89863	hypothetical prote	682	20	2	PN0459	basic eosinophil p
610	6	3.4	1228	2	C98219	proline dehydrogen	683	21	2	S03979	testosterone 15alp
611	6	3.4	1228	2	AG3067	proline dehydrogen	684	25	2	JA0171	dnak-type molecula
612	6	3.4	1234	2	C97606	hypothetical prote	685	28	2	A41763	extracellular prot
613	6	3.4	1234	2	AE2828	conserved hypothet	686	28	2	S12172	hypothetical prote

687	5	2.8	30	2	B84053	hypothetical prote	760	5	2.8	69	2	S45610	guanylate cyclase
688	5	2.8	33	2	PQ0150	dnak-type molecula	761	5	2.8	70	2	F84752	hypothetical prote
689	5	2.8	33	2	S34505	hypothetical prote	762	5	2.8	70	2	S43419	conserved hypotet
690	5	2.8	34	2	I50713	homeobox protein -	763	5	2.8	70	2	T43815	conserved hypotet
691	5	2.8	34	2	I51364	homeobox protein -	764	5	2.8	71	2	JH0349	T-cell receptor be
692	5	2.8	34	2	I38900	homeobox protein -	765	5	2.8	71	2	S21364	dnak-type molecula
693	5	2.8	34	2	I49145	homeobox protein -	766	5	2.8	71	2	T39627	hypothetical prote
694	5	2.8	35	4	S41911	hypothetical prote	767	5	2.8	71	2	I70114	arrestin-D - rat (
695	5	2.8	36	1	F2R21	photosystem II pro	768	5	2.8	72	2	C30517	Ig heavy chain pre
696	5	2.8	36	1	F2NT1	photosystem II pro	769	5	2.8	72	2	A71970	cbh3-type cytochro
697	5	2.8	36	1	F2KHD	photosystem II pro	770	5	2.8	72	2	C83219	hypothetical prote
698	5	2.8	36	2	I59139	dnak-type molecula	771	5	2.8	72	2	E81126	conserved hypotet
699	5	2.8	36	2	S60133	photosystem II pro	772	5	2.8	72	2	G85909	partial hypotet t
700	5	2.8	36	2	S01044	photosystem II pro	773	5	2.8	72	2	AE3444	hypothetical prote
701	5	2.8	36	2	S28767	photosystem II pro	774	5	2.8	72	2	F90221	lsu ribosomal prot
702	5	2.8	36	2	S07877	photosystem II pro	775	5	2.8	73	2	B64538	cbh3-type cytochro
703	5	2.8	36	2	A05019	photosystem II pro	776	5	2.8	74	1	CBK55E	cytochrome b559 co
704	5	2.8	36	2	JN0315	photosystem II pro	777	5	2.8	74	2	C96347	hypothetical prote
705	5	2.8	36	2	S58535	photosystem II pro	778	5	2.8	74	2	C81197	hypothetical prote
706	5	2.8	39	2	S75180	photosystem II pro	779	5	2.8	74	2	AH3104	conserved hypotet
707	5	2.8	39	2	AH2286	photosystem II pro	780	5	2.8	75	2	A81297	hypothetical prote
708	5	2.8	40	2	A40128	probable antigen 1	781	5	2.8	75	2	F70761	hypothetical prote
709	5	2.8	40	2	B97413	hypothetical prote	782	5	2.8	75	2	AF2565	hypothetical prote
710	5	2.8	41	2	E81565	hypothetical prote	783	5	2.8	76	2	D81094	hypothetical prote
711	5	2.8	43	2	S41388	protein 3a - human	784	5	2.8	76	2	C82511	hypothetical prote
712	5	2.8	44	2	PQ0181	naringenin-chalcon	785	5	2.8	76	2	G82739	hypothetical prote
713	5	2.8	44	2	H69676	phosphatase (Rapa)	786	5	2.8	77	2	G83542	hypothetical prote
714	5	2.8	44	2	S21028	m-factor precursor	787	5	2.8	77	2	S19575	epoxide hydrolase
715	5	2.8	46	2	A10624	probable DNA inver	788	5	2.8	78	2	T17801	hypothetical prote
716	5	2.8	47	2	F81118	hypothetical prote	789	5	2.8	78	2	T08483	kleA protein - Ent
717	5	2.8	48	2	D82534	hypothetical prote	790	5	2.8	78	2	S72611	hypothetical prote
718	5	2.8	49	2	JH0643	GTP-binding protel	791	5	2.8	78	2	S72971	hypothetical prote
719	5	2.8	49	2	S08452	hypothetical prote	792	5	2.8	78	2	C98182	hypothetical prote
720	5	2.8	50	2	G97562	hypothetical prote	793	5	2.8	79	2	F95210	conserved domain p
721	5	2.8	50	2	AG2783	hypothetical prote	794	5	2.8	80	2	C91287	hypothetical prote
722	5	2.8	51	2	F82759	hypothetical prote	795	5	2.8	80	2	B83097	hypothetical prote
723	5	2.8	51	2	F95849	hypothetical prote	796	5	2.8	80	2	A54059	protein kinase C (
724	5	2.8	52	2	T07434	photosystem II pro	797	5	2.8	80	2	A10707	cell division acti
725	5	2.8	52	2	B90462	hypothetical prote	798	5	2.8	81	1	CBYB55	cytochrome b559 co
726	5	2.8	52	2	T07269	hypothetical prote	799	5	2.8	81	2	F85101	hypothetical prote
727	5	2.8	54	2	S35697	leukocidin chain F	800	5	2.8	81	2	H70237	hypothetical prote
728	5	2.8	55	2	T43598	hypothetical prote	801	5	2.8	81	2	C83792	hypothetical prote
729	5	2.8	55	2	F56976	transfer complex p	802	5	2.8	82	1	S53882	cytochrome b559 co
730	5	2.8	55	2	A70070	hypothetical prote	803	5	2.8	82	2	E84301	hypothetical prote
731	5	2.8	55	2	AC2366	hypothetical prote	804	5	2.8	82	2	S39706	hypothetical prote
732	5	2.8	56	2	S65513	ribulose-bisphosph	805	5	2.8	82	2	B97702	hypothetical prote
733	5	2.8	56	2	S74630	hypothetical prote	806	5	2.8	82	2	AF2286	cytochrome b559 al
734	5	2.8	56	2	C69044	hypothetical prote	807	5	2.8	83	1	CBMT55	cytochrome b559 co
735	5	2.8	56	2	H98056	hypothetical prote	808	5	2.8	83	1	S00418	cytochrome b559 co
736	5	2.8	57	1	B64240	ribosomal protein	809	5	2.8	83	1	CBMT55	cytochrome b559 co
737	5	2.8	57	2	S26205	metallothionein -	810	5	2.8	83	1	CBMT5E	cytochrome b559 co
738	5	2.8	57	2	AC1756	hypothetical prote	811	5	2.8	83	1	A29956	cytochrome b559 co
739	5	2.8	58	2	I77464	lutetizing hormon	812	5	2.8	83	1	S03191	cytochrome b559 co
740	5	2.8	58	2	AD2652	hypothetical prote	813	5	2.8	83	1	S58568	cytochrome b559 co
741	5	2.8	59	2	B27606	Ig heavy chain v-a	814	5	2.8	83	1	A48310	cytochrome b559 co
742	5	2.8	60	2	B90490	hypothetical prote	815	5	2.8	83	1	S55789	cytochrome b559 co
743	5	2.8	60	2	T28340	ORF MSV178 hypothe	816	5	2.8	83	1	S01243	cytochrome b559 co
744	5	2.8	60	2	H95354	protein [Imported	817	5	2.8	83	1	CBLV55	cytochrome b559 co
745	5	2.8	61	2	PN0561	rod outer segment	818	5	2.8	83	1	IHKREG	high potential iro
746	5	2.8	61	2	D83274	conserved hypotet	819	5	2.8	83	2	T07257	cytochrome b559 co
747	5	2.8	64	2	S78309	hypothetical prote	820	5	2.8	83	2	T07475	cytochrome b559 co
748	5	2.8	64	2	S51273	hypothetical prote	821	5	2.8	83	2	B84174	hypothetical prote
749	5	2.8	64	2	T34792	probable transcript	822	5	2.8	84	2	S78337	cytochrome b559 co
750	5	2.8	64	2	S23540	hypothetical prote	823	5	2.8	84	2	S73312	cytochrome b559 al
751	5	2.8	65	2	C46448	hypothetical prote	824	5	2.8	84	2	S21365	dnak-type molecula
752	5	2.8	65	2	G81292	hypothetical prote	825	5	2.8	84	2	S33999	BP84R protein - Af
753	5	2.8	66	2	H81804	hypothetical prote	826	5	2.8	84	2	T18097	hypothetical prote
754	5	2.8	66	2	B81061	hypothetical prote	827	5	2.8	85	1	R3NT16	ribosomal protein
755	5	2.8	67	2	H90319	hypothetical prote	828	5	2.8	85	2	F75291	hypothetical prote
756	5	2.8	67	2	I54475	HLA-DNA-related sm	829	5	2.8	86	2	H90682	hypothetical prote
757	5	2.8	68	2	D95025	ribosomal protein	830	5	2.8	86	2	F64527	hypothetical prote
758	5	2.8	68	2	E97896	50S ribosomal prot	831	5	2.8	86	2	F71978	hypothetical prote
759	5	2.8	68	2	B03309	dnak-type molecula	832	5	2.8	86	2	C81000	hypothetical prote

833	5	2.8	86	2	S66089	conserved hypohet	906	5	2.8	99	2	AI2403	hypothetical prote
834	5	2.8	86	2	A35241	IgE Fc receptor ga	907	5	2.8	100	2	TJ01656	leucyl-tRNA synthe
835	5	2.8	86	2	A46461	T-cell receptor et	908	5	2.8	100	2	JO3032	ubiquinol-cytochro
836	5	2.8	86	2	S02118	IgE Fc receptor I.	909	5	2.8	100	2	AB2710	conserved hypohet
837	5	2.8	86	2	D85533	hypothetical prote	910	5	2.8	100	2	C97492	hypothetical prote
838	5	2.8	86	2	AG0130	conserved hypohet	911	5	2.8	100	2	B43998	hypothetical prote
839	5	2.8	86	2	F64766	yalb protein - Esc	912	5	2.8	100	2	D84081	hypothetical prote
840	5	2.8	87	2	S77861	ribosomal protein	913	5	2.8	100	2	AE2669	conserved hypohet
841	5	2.8	87	2	B86743	conserved hypohet	914	5	2.8	101	2	C31982	Ca2+-transporting
842	5	2.8	87	2	E90933	Cell division acti	915	5	2.8	101	2	JH0354	T-cell receptor be
843	5	2.8	87	2	A85782	Cell division acti	916	5	2.8	101	2	JC7579	ribosomal protein
844	5	2.8	87	2	C64932	Cell division acti	917	5	2.8	101	2	A29165	proteoglycan link
845	5	2.8	88	1	QQECC8	hypothetical prote	918	5	2.8	101	2	A05124	hypothetical prote
846	5	2.8	88	2	T07057	ribosomal protein	919	5	2.8	101	2	D75013	hypothetical prote
847	5	2.8	88	2	G05517	hypothetical prote	920	5	2.8	101	2	H97263	PTS system IIB com
848	5	2.8	88	2	A84166	hypothetical prote	921	5	2.8	101	2	AD0405	conserved hypohet
849	5	2.8	88	2	C82468	hypothetical prote	922	5	2.8	102	2	JA0170	dnak-type molecula
850	5	2.8	88	2	C83823	hypothetical prote	923	5	2.8	102	2	AF1661	phosphotransferase
851	5	2.8	89	2	A90339	hypothetical prote	924	5	2.8	102	2	AH1289	phosphotransferase
852	5	2.8	89	2	AC1576	carbon dioxide con	925	5	2.8	102	2	B24444	hypothetical rRNA
853	5	2.8	89	2	AH1222	carbon dioxide con	926	5	2.8	102	2	S28361	hypothetical prote
854	5	2.8	89	4	T46004	probable pseudogen	927	5	2.8	102	2	S69463	hypothetical prote
855	5	2.8	90	2	S05375	dnak-type molecula	928	5	2.8	102	2	D88428	protein R07E5.12 l
856	5	2.8	90	2	B91003	probable regulator	929	5	2.8	102	2	S44781	C30A5.4 protein -
857	5	2.8	90	2	G84991	hypothetical prote	930	5	2.8	102	2	A72739	probable DNA-direc
858	5	2.8	90	2	C83643	hypothetical prote	931	5	2.8	103	2	AD1003	conserved hypohet
859	5	2.8	90	2	A81245	hypothetical prote	932	5	2.8	103	2	G83241	conserved hypohet
860	5	2.8	90	2	H82023	probable lipoprote	933	5	2.8	103	2	I37384	FAS soluble protei
861	5	2.8	90	2	C69480	hypothetical prote	934	5	2.8	103	2	AG1563	hypothetical prote
862	5	2.8	90	2	H85847	hypothetical prote	935	5	2.8	103	2	AH1206	hypothetical prote
863	5	2.8	91	1	RGEBHD	mercuric resistanc	936	5	2.8	103	2	E75071	hypothetical prote
864	5	2.8	91	1	S09524	mercuric resistanc	937	5	2.8	104	1	CCFG	cytochrome c - bul
865	5	2.8	91	1	RGPSHA	mercuric resistanc	938	5	2.8	104	2	JN0643	naphthalene 1,2-di
866	5	2.8	91	2	I54404	T cell antigen rec	939	5	2.8	104	2	SO4719	ribosomal protein
867	5	2.8	91	2	S07895	hypothetical prote	940	5	2.8	104	2	G43721	recombination prot
868	5	2.8	91	2	JC7393	medaka-type gonado	941	5	2.8	104	2	S45400	probable membrane
869	5	2.8	91	2	S72298	hypothetical prote	942	5	2.8	104	2	A72649	hypothetical prote
870	5	2.8	92	2	A82141	integration host f	943	5	2.8	104	2	AE1430	B. subtilis Yulb p
871	5	2.8	92	2	A84397	50S ribosomal prot	944	5	2.8	104	2	S00562	homeotic protein E
872	5	2.8	92	2	F69390	LSU ribosomal prot	945	5	2.8	104	2	AE2131	hypothetical prote
873	5	2.8	92	2	PL0223	T-cell receptor be	946	5	2.8	105	2	I51851	parathyroid hormon
874	5	2.8	93	2	PL0224	T-cell receptor be	947	5	2.8	105	2	T10227	hypothetical prote
875	5	2.8	94	1	BCBOIA	S-100 protein alph	948	5	2.8	105	2	S70042	hypothetical prote
876	5	2.8	94	1	BCBOIA	S-100 protein alph	949	5	2.8	105	2	AB3256	hypothetical cytos
877	5	2.8	94	2	S03380	major fecal aller9	950	5	2.8	106	1	R8BY2B	acidic ribosomal p
878	5	2.8	94	2	A49832	DNA gyrase A - Sta	951	5	2.8	106	2	T16378	hypothetical prote
879	5	2.8	94	2	B49828	T-cell receptor be	952	5	2.8	106	2	T36973	hypothetical prote
880	5	2.8	94	2	C49828	T-cell receptor be	953	5	2.8	107	2	JH0352	T-cell receptor be
881	5	2.8	94	2	F86566	CT490 hypohet	954	5	2.8	107	2	S36262	Ig lambda chain v
882	5	2.8	94	2	T01217	hypohet	955	5	2.8	107	2	A69547	conserved hypohet
883	5	2.8	94	2	C72058	conserved hypohet	956	5	2.8	107	2	H82729	hypohet
884	5	2.8	94	2	D71261	hypohet	957	5	2.8	107	2	AI0282	probable DNA-bindi
885	5	2.8	95	2	A49027	T-cell receptor va	958	5	2.8	107	4	A60600	hypohet
886	5	2.8	95	2	H69767	hypohet	959	5	2.8	108	2	S36277	Ig lambda chain v
887	5	2.8	95	2	AC3625	hypohet	960	5	2.8	108	2	S36283	Ig lambda chain v
888	5	2.8	96	2	D46289	keratinocyte growt	961	5	2.8	108	2	S22892	T-cell receptor al
889	5	2.8	96	2	AF0741	conserved hypohet	962	5	2.8	108	2	S23204	retinol-binding pr
890	5	2.8	96	2	B90793	hypohet	963	5	2.8	108	2	E83765	hypohet
891	5	2.8	96	2	T31209	hypohet	964	5	2.8	108	2	AF0116	conserved hypohet
892	5	2.8	96	2	G85653	hypohet	965	5	2.8	108	2	AE0879	conserved hypohet
893	5	2.8	96	2	D95949	hypohet	966	5	2.8	108	2	T33603	hypohet
894	5	2.8	96	2	AF3175	attachment protein	967	5	2.8	109	2	H84139	PTS system, cellob
895	5	2.8	97	2	TJ0703	ferredoxin-Thiore	968	5	2.8	109	2	S58825	hypohet
896	5	2.8	97	2	B86345	Fl6P4.12 protein -	969	5	2.8	109	2	D70077	hypohet
897	5	2.8	98	2	S03497	T-cell receptor be	970	5	2.8	109	2	AE0849	hypohet
898	5	2.8	98	2	S17390	T-cell receptor be	971	5	2.8	109	2	S31950	penicillin-binding
899	5	2.8	98	2	PH0874	Ig heavy chain v r	972	5	2.8	110	1	R5BYA1	acidic ribosomal p
900	5	2.8	98	2	D96797	Sm-like protein (i	973	5	2.8	110	2	AG3382	quaternary ammoni
901	5	2.8	99	2	C60339	spap surface antig	974	5	2.8	110	2	A69902	hypohet
902	5	2.8	99	2	S07268	regulatory protein	975	5	2.8	110	2	E71004	hypohet
903	5	2.8	99	2	A70608	probable PE protei	976	5	2.8	110	2	AI2898	conserved hypohet
904	5	2.8	99	2	F86735	hypohet	977	5	2.8	111	2	G70521	hypohet
905	5	2.8	99	2	F82709	hypohet	978	5	2.8	111	2	T31580	hypohet

979 5 2.8 111 2 AD1313
980 5 2.8 111 2 A75322
981 5 2.8 112 2 JU0151
982 5 2.8 112 2 S25572
983 5 2.8 112 2 S69318
984 5 2.8 112 2 H71118
985 5 2.8 112 2 T47738
986 5 2.8 112 2 G70855
987 5 2.8 112 2 A87239
988 5 2.8 112 2 AE0496
989 5 2.8 113 1 CCCR6F
990 5 2.8 113 1 R6D0P1
991 5 2.8 113 2 JH0351
992 5 2.8 113 2 I38312
993 5 2.8 113 2 E27664
994 5 2.8 113 2 D27664
995 5 2.8 113 2 D27664
996 5 2.8 113 2 C70927
997 5 2.8 113 2 G87110
998 5 2.8 113 2 B40899
999 5 2.8 113 2 T44302
1000 5 2.8 113 2 C71074

ALIGNMENTS

RESULT 1

outer membrane protein P1 precursor, subtype 6U - Haemophilus influenzae (strain 8358, b)
C:Species: Haemophilus influenzae
C>Date: 28-Oct-1992 #sequence_revision 30-Jan-1993 #text_change 12-Dec-1997
C:Accession: B40183
R:Munson Jr., R.; Grass, S.; Elmhorn, M.; Bailey, C.; Newell, C.
A:Title: Comparative analysis of the structures of the outer membrane protein P1 genes of
Infect. Immun. 57, 3300-3305, 1989
A:Reference number: A40183; MUID:90035394; PMID:2572549
A:Accession: B40183
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <MUN>
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-450/Product: outer membrane protein P1 #status predicted <MAT>

Query Match 7.3% Score 13; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
|||||
Db 391 GATYKFTPNLSVD 403

RESULT 2

outer membrane protein P1 precursor, subtype 3L - Haemophilus influenzae (strain 1613, t)
C:Species: Haemophilus influenzae
C>Date: 28-Oct-1992 #sequence_revision 30-Jan-1993 #text_change 20-Aug-1999
C:Accession: A40183
R:Munson Jr., R.; Grass, S.; Elmhorn, M.; Bailey, C.; Newell, C.
A:Title: Comparative analysis of the structures of the outer membrane protein P1 genes of
Infect. Immun. 57, 3300-3305, 1989
A:Reference number: A40183; MUID:90035394; PMID:2572549
A:Accession: A40183
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <MUN>
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein

Query Match 7.3% Score 13; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
|||||
Db 391 GATYKFTPNLSVD 403

F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-450/Product: outer membrane protein P1 #status predicted <MAT>

Query Match 7.3% Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
|||||
Db 393 GATYKFTPNLSVD 405

RESULT 3

outer membrane protein P1 precursor - Haemophilus influenzae (type b)
C:Species: Haemophilus influenzae
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Aug-1999
C:Accession: A30510; A28787
R:Munson Jr., R.; Grass, S.
A:Title: Purification, cloning, and sequence of outer membrane protein P1 of Haemophilus
Infect. Immun. 56, 2235-2242, 1988
A:Reference number: A30510; MUID:88314258; PMID:2842261
A:Accession: A30510
A:Molecule type: DNA
A:Residues: 1-459 <MUN>
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-459/Product: outer membrane protein P1 #status predicted <MAT>

Query Match 7.3% Score 13; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
|||||
Db 397 GATYKFTPNLSVD 409

RESULT 4

outer membrane protein P1 precursor - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Aug-1999
C:Accession: F64065
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Klrkness, B.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gneim, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64065
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-459 <TIGR>
C:Superfamily: long-chain fatty acid transport protein fadL
C:Keywords: membrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-459/Product: outer membrane protein P1 #status predicted <MAT>

Query Match 7.3% Score 13; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
|||||
Db 397 GATYKFTPNLSVD 409

RESULT 5

G82406
long-chain fatty acid transport protein VCA0862 [imported] - Vibrio cholerae (strain N16)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82406
R:Heldeberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
i, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82406
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-412 <HEI>
A:Cross-references: GB:AE004414; GB:AE003853; NID:g958293; PIDN:AAF96760.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0862
A:Map position: 2
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 5.0%; Score 9; DB 2; Length 412;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 WELSGFHL 31
|||||||
Db 261 WELSGFHL 269

RESULT 6
H69448
heme d1 biosynthesis protein (nird) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69448
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69448
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-72 <KDE>
A:Cross-references: GB:AE000992; GB:AE000782; NID:g2689315; PIDN:AAB89654.1; PID:g264896

Query Match 4.5%; Score 8; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 YGLNLNVR 178
|||||||
Db 52 YGLNLNVR 59

RESULT 7
E90844
probable factor [imported] - Escherichia coli (strain O157:H7, substrain R1MD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90844
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90844
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-417 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA835148.1; PID:gl13361190; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECS1725

Query Match 4.5%; Score 8; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GLNLNVR 179
|||||||
Db 256 GLNLNVR 263

RESULT 8

D85702
hypothetical protein ychp [imported] - Escherichia coli (strain O157:H7, substrain E1
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85702
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Llm, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85702
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-417 <STO>
A:Cross-references: GB:AE005174; NID:gl2514938; PIDN:AAG56080.1; GSPDB:GN00145; UWGP
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ychp

Query Match 4.5%; Score 8; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GLNLNVR 179
|||||||
Db 256 GLNLNVR 263

RESULT 9

A64869
ychp protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: A64869
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A64869
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-417 <BLAT>

A:Cross-references: GB:AE000220; GB:U00096; NID:gl1787467; PIDN:AAC74304.1; PID:gl1787
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ychp; ycho

Query Match 4.5%; Score 8; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GLNLNVR 179
|||||||
Db 256 GLNLNVR 263

RESULT 10

F75107

abc transporter, ATP-binding protein PAB1929 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 17-Mar-2000

C:Accession: F75107

R:Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure

A:Reference number: A75001

A:Accession: F75107

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-633 <KAW>

A:CROSS-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49575.1; PID:el51547

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1929

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

F:438-623/Domain: ATP-binding cassette homology <ABC>

Query Match 4.5% Score 8; DB 2; Length 633;

Best Local Similarity 100.0%; Pred. No. 7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 KKKHFVET 149

Db 319 KKKHFVET 326

|||||

RESULT 11

D98068

degenerate transposase (orf1) [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: D98068

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: D98068

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-66 <KUP>

A:CROSS-references: GB:AE007317; PIDN:AAL00377.1; PID:gl5459240; GSPDB:GN00174

C:Genetics:

A:Gene: IS1381-truncation

Query Match

Best Local Similarity 3.9% Score 7; DB 2; Length 66;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 SKTHLSA 108

Db 40 SKTHLSA 46

|||||

RESULT 12

MYUY

myoglobin (validated) - yellowfin tuna

C:Species: Thunnus albacares (yellowfin tuna)

C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 15-Sep-2000

C:Accession: A02526

R:Watts, D.A.; Rice, R.H.; Brown, W.D.

J. Biol. Chem. 255, 10916-10924, 1980

A:Title: The primary structure of myoglobin from yellowfin tuna (Thunnus albacares).

A:Reference number: A02526; MUID:81046956; PMID:7430163

A:Accession: A02526

A:Molecule type: protein

A:Residues: 1-146 <WAT>

R:Birnbaum, G.I.; Evans, S.V.; Przybylska, M.; Rose, D.R.

submitted to the Brookhaven Protein Data Bank, May 1991

A:Reference number: A51306; PDB:1MWT

A:Contents: annotation: X-ray crystallography, 1.74 angstroms, residues 1-146

R:Lattman, E.E.; Nockolds, C.E.; Kretsinger, R.H.; Love, W.E.

J. Mol. Biol. 60, 271-277, 1971

A:Title: Structure of yellow fin tuna metmyoglobin at 6 angstroms resolution.

A:Reference number: A58455; MUID:72014155; PMID:5107328

A:Contents: annotation: X-ray crystallography, 6.0 angstroms

C:Function:

A:Description: binds molecular oxygen for intracellular storage and transport, primary

C:Superfamily: globin; globin homology

C:Keywords: acetylated amino end; chromoprotein; heme; iron; metalloprotein; muscle;

F:1-140/Domain: globin homology <GLB>

F:1/Modified site: acetylated amino end (Ala) #status experimental

F:59/Binding site: oxygen (His) (distal axial ligand) #status experimental

F:88/Binding site: heme iron (His) (proximal axial ligand) #status experimental

Query Match 3.9% Score 7; DB 1; Length 146;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 VEADYTT 164

Db 13 VEADYTT 19

|||||

RESULT 13

S00942

hypothetical protein 2 - Halobacterium sp. plasmid pHSB

C:Species: Halobacterium sp.

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 18-Nov-1994

C:Accession: S00942

R:Kagramanova, V.K.; Derckacheva, N.I.; Mankin, A.S.

Nucleic Acids Res. 16, 4158, 1988

A:Title: The complete nucleotide sequence of the arcaebacterial plasmid pHSB from Hal

A:Reference number: S00941; MUID:88234008; PMID:3287339

A:Accession: S00942

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-168 <KAG>

A:CROSS-references: EMBL:X07128

C:Genetics:

A:Genome: plasmid

Query Match 3.9% Score 7; DB 2; Length 168;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SRFAIGT 81

Db 26 SRFAIGT 32

|||||

RESULT 14

AB2849

hypothetical protein Atu2219 [imported] - Agrobacterium tumefaciens (strain C58, Dupo

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AB2849

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; MCCI

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AB2849

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-180 <KUP>

A:Cross-References: GB:AE008688; PIDN:AAL43208.1; PID:g17740689; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu2219
 A:Map position: circular chromosome

Query Match 3.9%; Score 7; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 TLRAGLA 96
 |||||
 DB 109 TLRAGLA 115

RESULT 15

F84748 probable AP2 domain transcription factor [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84748
 R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: F84748
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-References: GB:AE002093; NID:g1707016; PIDN:AAC69127.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g33710
 A:Map position: 2

Query Match 3.9%; Score 7; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 LAYDKAA 101
 |||||
 DB 107 LAYDKAA 113

RESULT 16

T02590 DNA binding protein EREBP-2 - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 21-Jul-2000
 C:Accession: T02590
 R:Ohme-Takagi, M.; Shinsli, H.
 Plant Cell 7, 173-182, 1995
 A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene response
 A:Reference number: Z14671; MUID:95276459; PMID:7756828
 A:Accession: T02590
 A>Status: preliminary; translated from GB/EMBL/DDBJ.
 A:Molecule type: mRNA
 A:Residues: 1-233 <OHM>
 A:Cross-References: EMBL:D38126; NID:g790362; PIDN:BAA07324.1; PID:g1208498
 A:Experimental source: strain BY4; tissue-type leaf

Query Match 3.9%; Score 7; DB 2; Length 233;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 LAYDKAA 101
 |||||
 DB 137 LAYDKAA 143

RESULT 17

CRM52

carbonate dehydratase (EC 4.2.1.1) II - mouse
 N:Alternate names: carbonic anhydrase II
 C:Species: Mus musculus (house mouse)
 C:Date: 25-Feb-1985 #sequence_revision 06-Feb-1995 #text_change 18-Jun-1999
 C:Accession: A23900; B23202; A01143; A20539; I51949
 R:Venta, P.J.; Montgomery, J.C.; Hewett-Emmet, D.; Wlebauer, K.; Tashian, R.E.
 J. Biol. Chem. 260, 12130-12135, 1985
 A:Title: Structure and exon to protein domain relationships of the mouse carbonic anhy
 A:Reference number: A23900; MUID:86008276; PMID:2995362
 A:Accession: A23900
 A:Molecule type: DNA
 A:Residues: 1-260 <VEN>
 A:Experimental source: strain YBR
 R:Venta, P.J.; Montgomery, C.; Hewett-Emmett, D.; Tashian, R.E.
 Biochim. Biophys. Acta 826, 195-201, 1985
 A:Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes
 A:Reference number: A90655; MUID:86077780; PMID:3000449
 A:Accession: B23202
 A:Molecule type: DNA
 A:Residues: 1-77 <VE2>
 A:Note: the authors translated the codon CAG for residue 39 as His
 R:Curtis, P.J.; Withers, E.; Demuth, D.; Watt, R.; Venta, P.J.; Tashian, R.E.
 Gene 25, 325-332, 1983
 A:Title: The nucleotide sequence and derived amino acid sequence of cDNA coding for m
 A:Reference number: A01143; MUID:84109569; PMID:6420240
 A:Accession: A01143
 A:Molecule type: mRNA
 A:Residues: 2-38, 'H', 40-260 <CUR>
 A:Cross-References: GB:K00811; GB:K00812; GB:M11830; NID:g192333; PIDN:AAA37356.1; PI
 A:Note: Initiator Met not shown
 R:Curtis, P.J.
 J. Biol. Chem. 258, 4459-4463, 1983
 A:Title: Cloning of mouse carbonic anhydrase mRNA and its induction in mouse erythro
 A:Reference number: A20539; MUID:83161023; PMID:6187736
 A:Accession: A20539
 A:Molecule type: mRNA
 A:Residues: 155-178; 214-240 <CU2>
 R:Venta, P.J.; Montgomery, J.C.; Wlebauer, K.; Hewett-Emmett, D.; Tashian, R.E.
 Ann. N. Y. Acad. Sci. 429, 309-323, 1984
 A:Title: Organization of the mouse and human carbonic anhydrase II genes.
 A:Reference number: I51949; MUID:84255152; PMID:6331255
 A:Accession: I51949
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 241-260 <RES>
 A:Cross-References: GB:M25944; NID:g199078; PIDN:AAA39505.1; PID:g199079
 C:Genetics:
 A:Gene: Car-2
 A:Introns: 12/1; 78/1; 117/3; 144/1; 169/3; 221/3
 C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase; zinc
 F:5-259/Domain: carbonic anhydrase homology <CAH>
 F:94,96,119/Binding site: zinc (His) #status predicted
 Query Match 3.9%; Score 7; DB 1; Length 260;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 97 YDKAASK 103
 |||||
 DB 51 YDKAASK 57

RESULT 18
 E86354 hypothetical protein F16L1.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86354
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86354
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <STO>
A:Cross-references: GB:AE005172; NID:99454531; PIDN:AAF87854.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 3.9% Score 7; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 36; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 95 LAYDKAA 101
| | | | |
Db 120 LAYDKAA 126

RESULT 19
B71877
hypothetical protein jhp0872 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: B71877
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923662
A:Accession: B71877
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <ARN>
A:Cross-references: GB:AE001517; GB:AE001439; NID:94155444; PIDN:AAD06452.1; PID:9415544
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0872

Query Match 3.9% Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 37; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 68 KEEYKDN 74
| | | | |
Db 221 KEEYKDN 227

RESULT 20
F64498
selenium donor protein - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 17-Mar-2000
C:Accession: F64498
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: F64498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <BLU>
A:Cross-references: GB:U67599; NID:91592201; PID:91592202; TIGR:MJ1591; PID:9
C:Genetics:

A:Map position: FOR1563343-1564269
A:Start codon: TTG
C:Superfamily: conserved hypothetical protein MJ0640

Query Match 3.9% Score 7; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 41; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 68 KEEYKDN 74
| | | | |
Db 265 KEEYKDN 271

RESULT 21
T35643
glycerol-3-phosphate dehydrogenase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-Jan-2000
C:Accession: T35643
R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z21585
A:Accession: T35643
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-336 <MUR>
A:Cross-references: EMBL:AL034447; PIDN:CAA22402.1; GSPDB:GN00070; SCOEDB:SC7A1.03
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC7A1.03
C:Superfamily: glycerol-3-phosphate dehydrogenase

Query Match 3.9% Score 7; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 44; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 90 TLRAGLA 96
| | | | |
Db 86 TLRAGLA 92

RESULT 22
B30338
late competence protein comGA - Bacillus subtilis
N:Alternate names: comG operon protein 1
C:Species: Bacillus subtilis
C:Date: 27-Feb-1990 #sequence_revision 03-Aug-1992 #text_change 24-Oct-2000
C:Accession: B30338; B69603
R:Albano, M.; Breitling, R.; Dubnau, D.A.
J. Bacteriol. 171, 5386-5404, 1989
A:Title: Nucleotide sequence and genetic organization of the *Bacillus subtilis* comG o
A:Reference number: A30338; MUID:90008773; PMID:2507524
A:Accession: B30338
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <ALB>
A:Cross-references: GB:M29691; GB:M22854; NID:g1119199; PIDN:AAA83367.1; PID:g142706
A:Note: the authors translated the codon GAT for residue 182 as Glu
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caidwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabet, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumsteln, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69603
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-356 <KUN>
A:Cross-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CABL4404.1; PID:g2634907
A:Experimental source: strain 168
C:Genetics:
A:Gene: comGA
C:Superfamily: twitching motility protein pili

Query Match 3.9%; Score 7; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GSLTLKL 19
|||||
Db 77 GSLTLKL 83

RESULT 23
C84179
aminopeptidase homolog [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84179
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84179
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <STO>
A:Cross-references: GB:AE004437; NID:g10579833; PIDN:AAG18799.1; GSPDB:GN00138
C:Genetics:
A:Gene: pepB1

Query Match 3.9%; Score 7; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 EADYTTK 165
|||||
Db 70 EADYTTK 76

RESULT 24
A70795
hypothetical protein RV3712 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70795
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70795
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-413 <COL>
A:Cross-references: GB:AL123456; NID:g3261559; PIDN:CAAL8034.1; PID:e136457
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV3712

Query Match 3.9%; Score 7; DB 2; Length 413;

Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 TLRAGLA 96
|||||
Db 148 TLRAGLA 154

RESULT 25
AH0334
probable long-chain fatty acid transport protein fadL [imported] - Yersinia pestis (s
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0334
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, J
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, C
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barre
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92983.1; PID:g15980722; GSPDB:GN00175
C:Genetics:
A:Gene: fadL
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 3.9%; Score 7; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 QWAIHYS 40
|||||
Db 281 QWAIHYS 287

RESULT 26
AE0805
long-chain fatty acid transport protein precursor STY2623 [imported] - Salmonella en
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0805
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Par
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s.
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0805
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07623.1; PID:g16503614; GSPDB:GN00176
C:Genetics:
A:Gene: STY2623
C:Superfamily: long-chain fatty acid transport protein fadL

Query Match 3.9%; Score 7; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 QWAIHYS 40
|||||
Db 302 QWAIHYS 308

RESULT 27
F65007
long-chain fatty acid transport protein fadL precursor - Escherichia coli (strain K-

N:Alternate names: FLP protein
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: F65007; A39126; S02829
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
C:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F65007
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-448 <BLAT>
A:Cross-references: GB:AE000323; GB:U00096; NID:gl788684; PIDN:AACT5404.1; PID:g1788686;
A:Experimental source: strain K-12, substrain MG1655
R:Black, P.N.
J. Bacteriol. 173, 435-442, 1991
A:Title: Primary sequence of the Escherichia coli fadL gene encoding an outer membrane protein
A:Reference number: A39126; MUID:91100327; PMID:1987139
A:Accession: A39126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177, 'V', 179-287, 'R', 289-448 <BLA>
A:Cross-references: GB:M60607; GB:M37714; NID:g145909; PIDN:AAA64433.1; PID:g145910
R:Salz, B.; Ghosh, C.R.; Vu, L.; Nunn, W.D.
Mol. Microbiol. 2, 363-370, 1988
A:Title: Nucleotide sequencing and expression of the fadL gene involved in long-chain fatty acid metabolism
A:Reference number: S02829; MUID:98288050; PMID:2840553
A:Accession: S02829
A:Molecule type: DNA
A:Residues: 66-177, 'R', 179, 'p', 181-287, 'R', 289-448 <SAI>
A:Cross-references: EMBL:Y00552; NID:g41371; PIDN:CAA68630.1; PID:g41372
A:Note: the authors translated the codon CTG for residue 162 as Glu
C:Genetics:
A:Gene: fadL
C:Superfamily: long-chain fatty acid transport protein fadL
Query Match 3.9%; Score 7; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 QWAIHYS 40
| | | | | | | |
Db 313 QWAIHYS 319
RESULT 28
long-chain fatty acid transport protein FadL ECS3227 [imported] - Escherichia coli (strain C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C91032
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic analysis
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C91032
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036650.1; PID:g13362697; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS3227
C:Superfamily: long-chain fatty acid transport protein fadL
Query Match 3.9%; Score 7; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 QWAIHYS 40
| | | | | | | |
Db 313 QWAIHYS 319
RESULT 28
long-chain fatty acid transport protein FadL ECS3227 [imported] - Escherichia coli (strain C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C91032
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic analysis
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C91032
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036650.1; PID:g13362697; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS3227
C:Superfamily: long-chain fatty acid transport protein fadL
Query Match 3.9%; Score 7; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 QWAIHYS 40
| | | | | | | |
Db 313 QWAIHYS 319
RESULT 29
long-chain fatty acid transport protein FadL - Escherichia coli (strain O157:H7, subspecies: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C:Accession: D85876
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May, J.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimaanta, E.; Potamousis, K.; Apodaca, N.; Rasmussen, S.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.; Tomb, J.F.; White, O.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee, J.D.; Keiley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <STO>
A:Cross-references: GB:AF005174; NID:g12516711; PIDN:AAG57472.1; GSPDB:GN00145; UWGP: A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fadL
C:Superfamily: long-chain fatty acid transport protein fadL
Query Match 3.9%; Score 7; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 QWAIHYS 40
| | | | | | | |
Db 313 QWAIHYS 319
RESULT 30
probable outer membrane protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: A71859
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovls, G.F. Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: A71859
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <ARN>
A:Cross-references: GB:AE001530; GB:AE001439; NID:g4155502; PIDN:AAD06606.1; PID:g415
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp1022
Query Match 3.9%; Score 7; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 12 KGSJTLK 18
| | | | | | | |
Db 283 KGSJTLK 289
RESULT 31
F64564
hypothetical protein HP0358 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: F64564
R:Tomb, J.F.; White, O.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee, J.D.; Keiley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <STO>
A:Cross-references: GB:AF005174; NID:g12516711; PIDN:AAG57472.1; GSPDB:GN00145; UWGP: A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fadL
C:Superfamily: long-chain fatty acid transport protein fadL
Query Match 3.9%; Score 7; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 QWAIHYS 40
| | | | | | | |
Db 313 QWAIHYS 319

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: F64564
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-511 <TOM>
A;Cross-references: GB:AE000552; GB:AE000511; NID:g2313451; PIDN:AAD07427.1; PID:g231346

Query Match 3.9%; Score 7; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 KGSUTLK 18
|||||||
Db 283 KGSUTLK 289

RESULT 32
S77572
oligopeptide transport protein homolog ophA - Agrobacterium tumefaciens plasmid pTIR10
N;Alternate names: oligopeptide permease homolog ophA
C;Species: Agrobacterium tumefaciens
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 17-Mar-2000
C;Accession: S77572
R;Fuqua, C.; Winans, S.C.
Mol. Microbiol. 20, 1199-1210, 1996
A;Title: Localization of OcrR-activated and TraR-activated promoters that express two AB
A;Reference number: S77571; MUID:96405643; PMID:8809772
A;Accession: S77572
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-519 <FUQ>
A;Cross-references: EMBL:U48718; NID:g1215729; PIDN:AAC44508.1; PID:g1215731
A;Experimental source: strain R10
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C;Genetics:
A;Gene: ophA
A;Genome: plasmid pTIR10
C;Superfamily: dipeptide transport protein

Query Match 3.9%; Score 7; DB 2; Length 519;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 TLRAGLA 96
|||||||
Db 89 TLRAGLA 95

RESULT 33
T08541
hypothetical protein F27B13.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C;Accession: T08541
R;Bevan, M.; Zimmermann, W.; Gruenelsen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A;Reference number: z16442
A;Accession: T08541
A;Molecule type: DNA
A;Residues: 1-526 <BEV>
A;Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.40
A;Experimental source: cultivar Columbia; BAC clone F27B13
C;Genetics:
A;Gene: ATSP:F27B13.40
A;Map position: 4
A;Introns: 421/3

Query Match 3.9%; Score 7; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 141 GKXKHEV 147
|||||||
Db 17 GKXKHEV 23

RESULT 34

AIO359
iron(III)-transport system permease sfuB [imported] - *Yersinia pestis* (strain CO92)
C;Species: *Yersinia pestis*
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C;Accession: AIO359
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, J.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Skellern, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barral
Nature 413, 523-527, 2001
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AIO359
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-528 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92204.1; PID:g15980916; GSPDB:GN00175
C;Genetics:
A;Gene: sfuB
C;Superfamily: sfuB protein

Query Match 3.9%; Score 7; DB 2; Length 528;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 GLLIVEA 160
|||||||
Db 250 GLLIVEA 256

RESULT 35

B45808
B-lymphocyte antigen CD19 precursor - mouse
C;Species: *Mus musculus* (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 02-Aug-1996 #text_change 02-Aug-2002
C;Accession: B45808; I49579; I49580
R;Tedder, T.F.; Isaacs, C.W.
J. Immunol. 143, 712-717, 1989
A;Title: Isolation of cDNAs encoding the CD19 antigen of human and mouse B lymphocytes
A;Reference number: A45808; MUID:89292699; PMID:2472450
A;Accession: B45808
A;Molecule type: mRNA
A;Residues: 302-547 <TED>
A;Cross-references: GB:M28240; NID:g192827; PIDN:AAA74753.1; PID:g192828
A;Note: the authors translated the codon GAC for residue 520 as Tyr
R;Zhou, L.J.; Ord, D.C.; Hughes, A.L.; Tedder, T.F.
J. Immunol. 147, 1424-1432, 1991
A;Title: Structure and domain organization of the CD19 antigen of human, mouse, and
A;Reference number: I48142; MUID:91332460; PMID:1714482
A;Accession: I49579
A;Status: translation not shown; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-547 <RES>
A;Cross-references: GB:M62542; NID:g192462; PIDN:AAA37388.1; PID:g192463
R;Zhou, L.J.; Ord, D.C.; Omori, S.A.; Tedder, T.F.
Immunogenetics 35, 102-111, 1992
A;Title: Structure of the genes encoding the CD19 antigen of human and mouse B lymphocytes
A;Reference number: I49580; MUID:92137946; PMID:1370948
A;Accession: I49580
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-206, 'A', 208-547 <RE2>
A;Cross-references: GB:M62553; NID:g192468; PIDN:AAA37390.1; PID:g192470
C;Genetics:
A;Gene: CD19

A;Introns: 30/1; 119/1; 184/1; 277/1; 314/1; 330/2; 360/1; 394/1; 428/1; 451/1; 470/
C;Superfamily: human B-cell antigen CD19; immunoglobulin homology
C;Keywords: B-cell; glycoprotein; surface antigen; transmembrane protein

F:31-99/Domain: immunoglobulin homology <IMM1>
F:190-261/Domain: immunoglobulin homology <IMM2>

Query Match 3.9%; Score 7; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 57 QDGSYGE 63
|||||||
Db 427 QDGSYGE 433

RESULT 36

Fiber protein - human adenovirus 31

C:Species: Mastadenovirus h31 (human adenovirus 31)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C:Accession: S39297
R:Pring-Akerblom, P.; Adrian, T.
submitted to the EMBL Data Library, November 1993.
A:Reference number: S39296
A:Accession: S39297
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-556 <PRI>
A:Cross-references: EMBL:X76548; NID:9434907; PIDN:CAA54050.1; PID:g434908
C:Superfamily: adenovirus fiber protein

Query Match 3.9%; Score 7; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 GSLTLKL 19
|||||||
Db 66 GSLTLKL 72

RESULT 37

A44441

B-cell antigen CD19 precursor - human

N:Alternate names: B-cell-specific glycoprotein CD19; earliest B cell lineage differentials
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 02-Aug-1996 #text_change 02-Aug-2002
C:Accession: A44441; J00074; A45808; I84731
R:Kozmik, Z.; Wang, S.; Doerfler, P.; Adams, B.; Busslinger, M.
Mol. Cell. Biol. 12, 2662-2672, 1992
A:Title: The promoter of the CD19 gene is a target for the B-cell-specific transcription
A:Reference number: A44441; MUID:92269839; PMID:1375324
A:Accession: A44441
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-556 <KOZ>
A:Cross-references: GB:M84371; NID:g901822; PIDN:AAA69966.1; PID:g901823
R:Stamenkovic, I.; Seed, B.
J. Exp. Med. 168, 1205-1210, 1988
A:Title: CD19, the earliest differentiation antigen of the B cell lineage, bears three e
A:Reference number: J00074; MUID:89010509; PMID:2459292
A:Accession: J00074
A:Molecule type: mRNA
A:Residues: 1-79, 'I', 81-395, 'RREWAQKKRGRAMRLTVRRTPSSMRRTPTTLGRTSSPRMAAATRLRSPWVLRMT
A:Cross-references: GB:M21057; NID:g178666; PIDN:AAA35533.1; PID:g178667
A>Note: the extracellular domain is organized in 19-like domains
A>Note: the region beyond residue 395 differs from all other sequence reports in this en
Epstein-Barr virus BSLF1 protein (see entry PIR1:Q08E15)
R:Tedder, T.F.; Isaacs, C.M.
J. Immunol. 143, 712-717, 1989
A:Title: Isolation of cDNAs encoding the CD19 antigen of human and mouse B lymphocytes.
A:Reference number: A45808; MUID:89292699; PMID:2472450
A:Accession: A45808
A:Molecule type: mRNA
A:Residues: 1-513, 'H', 515-556 <TED>
A:Cross-references: GB:M28170; NID:g862622; PIDN:AAA68490.1; PID:g387662
R:Zhou, L.J.; Ord, D.C.; Omori, S.A.; Tedder, T.F.

Immunogenetics 35, 102-111, 1992

A:Title: Structure of the genes encoding the CD19 antigen of human and mouse B lympho
A:Reference number: I49580; MUID:92137946; PMID:1370948
A:Accession: I84731
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-556 <RES>
A:Cross-references: GB:M62550; NID:g180032; PID:g180034
C:Genetics:

A:Gene: GDB:CD19
A:Cross-references: GDB:127605; OMIM:107265

A:Map position: 16p11.2-16p11.2
A:Introns: 30/1; 119/1; 187/1; 279/1; 316/1; 332/2; 362/1; 400/1; 435/1; 458/1; 477/1
C:Superfamily: human B-cell antigen CD19; immunoglobulin homology
C:Keywords: B-cell; glycoprotein; surface antigen; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-556/Product: B-cell specific glycoprotein CD19 #status predicted <MAT>
F:21-291/Domain: extracellular #status predicted <EXT>
F:193-263/Domain: immunoglobulin homology <IMM>
F:292-319/Domain: transmembrane #status predicted <TMD>
F:310-556/Domain: intracellular #status predicted <INT>
F:86,125,138,181,265/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.9%; Score 7; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 57 QDGSYGE 63
|||||||
Db 434 QDGSYGE 440

RESULT 38

AD1666

DNA ligase homolog lin1870 [imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1666
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,
J.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1666
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-671 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97100.1; PID:g16414371; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:

A:Gene: lin1870
C:Superfamily: polydeoxyribonucleotide synthase (NAD+)

Query Match 3.9%; Score 7; DB 2; Length 671;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 92 RAGLAYD 98
|||||||
Db 270 RAGLAYD 276

RESULT 39

T47474

hypothetical protein F18N11.100 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47474
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000

A:Reference number: 224467
A:Accession: T47474
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-693 <JOR>
A:Cross-references: EMBL:AL132953
A:Experimental source: cultivar Columbia; BAC clone F18N11
C:Genetics:
A:Map position: 3
A:Introns: 18/3; 630/3
A:Note: F18N11.100

Query Match 3.9%; Score 7; DB 2; Length 693;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRGKKKH 145
Db 319 LRGKKKH 325
|||||||

RESULT 40

T02196
hypothetical protein At2g46910 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F14M4.26
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
A:Accession: T02196; G84908
R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Shen, M.; Roun-
submitted to the EMBL Data Library, September 1998
A:Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.
A:Reference number: 214609
A:Accession: T02196
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-734 <ROU>
A:Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522947
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-734 <STO>
A:Cross-references: GB:AE002093; NID:g3522947; PIDN:AAC34229.1; GSPDB:GN00139
C:Genetics:
A:Gene: F14M4.26; At2g46910
A:Map position: 2
A:Introns: 138/2; 167/3; 214/2; 259/3; 369/2; 406/3; 467/2; 513/3; 535/3; 547/3; 560/1;
C:Superfamily: Arabidopsis thaliana hypothetical protein F14M4.26

Query Match 3.9%; Score 7; DB 2; Length 734;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 KGLLVE 159
Db 223 KGLLVE 229
|||||||

Search completed: May 12, 2003, 09:49:10
Job time : 53 secs

GenCore version 5.1.4.p5.4578
Copyright, (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:42:05 ; Search time 12 seconds
(without alignments)
618.689 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 179

Sequence: 1 QHNGVLGPYIGKSLTKLP.....ADYTTKATANLYGLNLNRYF 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Swissprot_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	7.3	459	1 OM11_HAEIN	P43838 haemophilus
2	13	7.3	459	1 OM12_HAEIN	P10841 haemophilus
3	8	4.5	417	1 YCHO_ECOLI	P39165 escherichia
4	7	3.9	146	1 MYG_GDRGI	Q9den9 goblionototh
5	7	3.9	146	1 MYG_THUAL	P02205 thunnus alb
6	7	3.9	146	1 MYG_THUTH	Q9dd47 thunnus thy
7	7	3.9	259	1 CAH2_MOUSE	P00920 mus musculus
8	7	3.9	336	1 GPDA_STRCO	Q92850 streptomyce
9	7	3.9	356	1 CMGA_BACSU	P25953 bacillus su
10	7	3.9	415	1 HN3B_ORYLA	O42097 oryzias lat
11	7	3.9	428	1 OM47_PASMU	P80603 pasteurella
12	7	3.9	448	1 FADL_ECOLI	P10384 escherichia
13	7	3.9	466	1 AMID_PSEPU	O69768 pseudomonas
14	7	3.9	547	1 CD19_MOUSE	P25918 mus musculus
15	7	3.9	556	1 CD19_HUMAN	P15391 homo sapien
16	7	3.9	556	1 FIBP_ADE31	P36848 human adeno
17	7	3.9	859	1 RPA2_METJA	Q58446 methanococc
18	7	3.9	865	1 HTRE_ECOLI	P33129 escherichia
19	7	3.9	977	1 A2AL_MOUSE	P17426 m adapter-r
20	7	3.9	1641	1 YC52_XYLFA	Q9pdx7 xylella fas
21	6	3.4	52	1 ZN32_HUMAN	P17041 homo sapien
22	6	3.4	77	1 RS20_LACLA	Q9ceus5 lactococcus
23	6	3.4	86	1 RS20_MYCLE	P33132 mycobacteri
24	6	3.4	86	1 RS20_MYCTU	P71731 mycobacteri
25	6	3.4	94	1 RL35_PIG	Q29361 sus scrofa
26	6	3.4	97	1 Y004_BHP1	P51703 bacterioph
27	6	3.4	104	1 RK21_PORPU	P51209 porphyra pu
28	6	3.4	120	1 YHFY_ECOLI	P45551 escherichia
29	6	3.4	122	1 RL35_HUMAN	P42766 homo sapien
30	6	3.4	122	1 RL35_RAT	P17078 rattus norv
31	6	3.4	125	1 MSP1_GLORO	P53021 globodera r
32	6	3.4	125	1 MSP2_GLORO	P53022 globodera r
33	6	3.4	125	1 MSP3_GLORO	P53023 globodera r

34	6	3.4	139	1 FXSA_SBRMA	P37148 serratia ma
35	6	3.4	146	1 MYG_PSEGE	Q9depl pseudochaen
36	6	3.4	153	1 YXIK_BACSU	P42302 bacillus su
37	6	3.4	180	1 GRPE_STRMU	O06941 streptococc
38	6	3.4	184	1 ITC_ALOMA	P35812 alocasia ma
39	6	3.4	185	1 PTH_BACHD	Q9kgj3 bacillus ha
40	6	3.4	186	1 DYR_HSVS7	P27421 herpesvirus
41	6	3.4	187	1 PABA_ECOLI	P00903 escherichia
42	6	3.4	187	1 PABA_SALTY	P06193 salmonella
43	6	3.4	191	1 SC21_RICPR	Q928b4 rickettsia
44	6	3.4	194	1 CLP1_CYPAP	Q36863 cyanophora
45	6	3.4	199	1 SC21_RICCN	Q94jm5 rickettsia
46	6	3.4	202	1 RR41_CVACA	Q9tli7 cyanidium c
47	6	3.4	206	1 Y615_SYNY3	P52876 synecocyst
48	6	3.4	210	1 RL4_THETH	P49665 thermus the
49	6	3.4	213	1 DYR_HSVSC	P22573 herpesvirus
50	6	3.4	216	1 ANSP_EIMTE	P15744 eimeria ten
51	6	3.4	224	1 UPP_SYNY3	P72753 synecocyst
52	6	3.4	224	1 RNT_VIBCH	Q9k497 vibrio chol
53	6	3.4	224	1 UPP_TOBAC	P93394 nicotiana t
54	6	3.4	233	1 ATTB_HVACE	P01512 hyalophora
55	6	3.4	241	1 LEG3_RABIT	P47845 oryctolagus
56	6	3.4	243	1 YC23_CYPAP	P31605 cyanophora
57	6	3.4	244	1 LEG3_CRILQ	P47953 cricetulus
58	6	3.4	247	1 VATD_BOVIN	P39942 bos taurus
59	6	3.4	247	1 VATD_HUMAN	Q9y5k8 homo sapien
60	6	3.4	247	1 VATD_MOUSE	P57746 mus musculu
61	6	3.4	247	1 VATD_RABIT	Q97755 oryctolagus
62	6	3.4	250	1 H11_DROVI	Q24704 drosophila
63	6	3.4	250	1 H12_DROVI	Q94555 drosophila
64	6	3.4	250	1 H13_DROVI	Q94772 drosophila
65	6	3.4	250	1 VATD_SUBDO	P57747 suberites d
66	6	3.4	259	1 Y348_CHLPN	Q92815 chlamydia p
67	6	3.4	261	1 LEG3_RAT	P08699 rattus norv
68	6	3.4	262	1 TRUA_PYRHO	O58941 pyrococcus
69	6	3.4	263	1 KDGK_ECOLI	P76268 escherichia
70	6	3.4	263	1 LEG3_MOUSE	P16110 mus musculu
71	6	3.4	267	1 PYRF_KLUMA	P41769 kluyveromyc
72	6	3.4	267	1 RM09_HUMAN	Q9byd2 homo sapien
73	6	3.4	267	1 YD24_SCHPO	O13711 schizosacch
74	6	3.4	282	1 TAUD_ECOLI	P37610 escherichia
75	6	3.4	282	1 YQ05_BACAN	Q9rn27 bacillus an
76	6	3.4	286	1 KDSA_FUSNN	O8re91 fusobacteri
77	6	3.4	288	1 A41_LEIDO	P55905 leishmania
78	6	3.4	289	1 NOLT_RHIFR	P33209 rhizobium f
79	6	3.4	289	1 NOLT_RHISN	P55714 rhizobium s
80	6	3.4	290	1 TODE_PSEPU	P13453 pseudomonas
81	6	3.4	290	1 VG33_HSVIL	Q00118 ictaluriid h
82	6	3.4	290	1 YABG_BACSU	P37548 bacillus su
83	6	3.4	294	1 EFTS_MYCPU	Q98q37 mycoplasma
84	6	3.4	295	1 LEG3_CANFA	P38486 canis famil
85	6	3.4	297	1 RRPB_RABVC	P15198 rabies viru
86	6	3.4	297	1 RRPB_RABVC	P22363 rabies viru
87	6	3.4	297	1 SG54_DROME	Q00725 drosophila
88	6	3.4	299	1 T2B1_BACSU	O68557 bacillus su
89	6	3.4	302	1 HRPS_PSESY	P37931 pseudomonas
90	6	3.4	302	1 OXYR_ERWCA	P71318 erwinia car
91	6	3.4	305	1 KDGK_ERWCH	P37728 erwinia chr
92	6	3.4	305	1 MEPL_SOYBN	P29136 glycine max
93	6	3.4	305	1 OXYR_ERWCH	Q9x725 erwinia chr
94	6	3.4	305	1 YN45_DEIRA	Q9rry7 delinococcus
95	6	3.4	308	1 VG10_BPB03	O37891 bacterioph
96	6	3.4	313	1 YK89_ARCFU	O28190 archaeoglob
97	6	3.4	317	1 PIP_SERMA	O32449 serratia ma
98	6	3.4	319	1 PYRB_MYCTU	P71808 mycobacteri
99	6	3.4	319	1 Y319_MYAXA	P96554 myxococcus
100	6	3.4	321	1 PYRB_MYCLE	Q9ctr5 mycobacteri
101	6	3.4	327	1 Y745_HELPJ	Q92198 helicobacte
102	6	3.4	327	1 Y745_HELPY	O25441 helicobacte
103	6	3.4	329	1 XYL1_KLUJA	P49378 kluyveromyc
104	6	3.4	335	1 HYPB_RHOCA	P26410 rhodobacter
105	6	3.4	345	1 YE6A_HAEIN	O86241 haemophilus
106	6	3.4	347	1 RTCA_RALSO	Q8y2v6 ralstonia s

107	1	YBGO_ECOLI	353	6	3.4	P75748	escherichla	180	6	3.4	608	1	HSCA_BUCAP	O51883	buchnera ap
108	1	T2H2_HAEPRA	358	6	3.4	P36433	haemophilus	181	6	3.4	611	1	HSCA_BUCAI	P57660	buchnera ap
109	1	RECF_DEIRA	359	6	3.4	Q9ve0	deinococcus	182	6	3.4	614	1	NCO4_HUMAN	Q13772	homo sapien
110	1	ARG1_XENLA	360	6	3.4	Q91553	xenopus lae	183	6	3.4	621	1	GIDA_HELPJ	O92m19	helicobacte
111	1	ARG2_XENLA	361	6	3.4	Q91554	xenopus lae	184	6	3.4	621	1	GIDA_HELPJ	P56138	helicobacte
112	1	ARG3_XENLA	360	6	3.4	Q91555	xenopus lae	185	6	3.4	622	1	AMAI_PLAF8	P50492	plasmodium
113	1	DNAJ_BORBU	364	6	3.4	P28616	borrelia bu	186	6	3.4	622	1	AMAI_PLAF8	P50489	plasmodium
114	1	P37_MYCCE	368	6	3.4	Q49410	mycoplasma	187	6	3.4	622	1	AMAI_PLAF8	P22621	plasmodium
115	1	DP3B_TREPA	371	6	3.4	O83048	treponema p	188	6	3.4	622	1	AMAI_PLAF8	P50490	plasmodium
116	1	RLPA_ANASP	371	6	3.4	O8vz24	anabena sp	189	6	3.4	622	1	AMAI_PLAF8	P50491	plasmodium
117	1	SWE_HAEIN	373	6	3.4	P43862	haemophilus	190	6	3.4	644	1	DXS_CHLPN	Q92639	chlamydia p
118	1	SBP_CRYUA	374	6	3.4	P18632	cryptomeria	191	6	3.4	644	1	LONI_METTH	Q26878	methanobact
119	1	YB12_MYCTU	378	6	3.4	Q10400	mycobacteri	192	6	3.4	658	1	CPT2_HUMAN	P23786	homo sapien
120	1	GC14_YEAST	383	1	3.4	P46959	saccharomyc	193	6	3.4	658	1	CPT2_HUMAN	P52825	mus musculu
121	1	Y181_ARCFU	390	1	3.4	Q30056	archaeoglob	194	6	3.4	659	1	AMIA_STRPN	P18791	streptococc
122	1	CEMA_NEPOL	392	1	3.4	Q9tk22	nephroselm	195	6	3.4	661	1	CL80_HUMAN	Q99467	homo sapien
123	1	EFTU_CHLVI	393	1	3.4	P42473	chlorobium	196	6	3.4	662	1	SL51_RABIT	P11170	oryctolagus
124	1	LOLC_HAEIN	394	1	3.4	P44252	haemophilus	197	6	3.4	676	1	LSHR_CALJA	Q02721	callithrix
125	1	YIE2_YEAST	394	1	3.4	P40530	saccharomyc	198	6	3.4	686	1	TGM2_MOUSE	P21981	mus musculu
126	1	DNAJ_STRCO	399	1	3.4	P40170	streptomyc	199	6	3.4	689	1	TGM2_MOUSE	P08587	cavia cutle
127	1	DRG_CAUCR	399	1	3.4	Q9a709	caulobacter	200	6	3.4	693	1	TGM3_HUMAN	Q08188	homo sapien
128	1	AROC_MYCTU	401	1	3.4	P95013	mycobacteri	201	6	3.4	693	1	TGM3_MOUSE	Q08189	mus musculu
129	1	PI52_MOUSE	405	1	3.4	O70172	mus musculu	202	6	3.4	698	1	ALT_BPT2	Q38424	bacterioph
130	1	G6PI_CAMJE	406	1	3.4	Q9pmd4	campylobact	203	6	3.4	698	1	ALT_BPT6	Q38433	bacterioph
131	1	PI52_HUMAN	406	1	3.4	P48426	homo sapien	204	6	3.4	702	1	DDX4_MOUSE	Q61496	mus musculu
132	1	PI53_HUMAN	406	1	3.4	P53807	homo sapien	205	6	3.4	702	1	DDX4_MOUSE	Q9hsc3	halobacteri
133	1	ODPA_KLULA	408	1	3.4	O13366	kluyveromyc	206	6	3.4	713	1	DDX4_RAT	Q64060	rattus norv
134	1	IRX5_HUMAN	417	1	3.4	P78411	homo sapien	207	6	3.4	722	1	FLID_TREPA	Q83842	treponema p
135	1	GSA_METTH	420	1	3.4	O26330	methanobact	208	6	3.4	724	1	DDX4_HUMAN	Q9nq10	homo sapien
136	1	KICR_MOUSE	422	1	3.4	P05784	mus musculu	209	6	3.4	724	1	YMF6_YEAST	Q05029	saccharomyc
137	1	AP2_ARATH	432	1	3.4	P47927	arabidopsis	210	6	3.4	759	1	MCMS_CAEEL	Q21902	caenorhabdi
138	1	PEPW_LACDL	437	1	3.4	P94868	lactobacilli	211	6	3.4	759	1	Y938_MYCTU	P71571	mycobacteri
139	1	PNK1_HUMAN	446	1	3.4	O95544	homo sapien	212	6	3.4	763	1	PEPX_LACLA	Q9ce01	lactococcus
140	1	P2X5_RAT	455	1	3.4	P51578	rattus norv	213	6	3.4	763	1	PEPX_LACLA	P22346	lactococcus
141	1	VATB_METBA	460	1	3.4	P22663	methanosarc	214	6	3.4	767	1	YNY7_YEAST	P53852	saccharomyc
142	1	VATB_METMA	460	1	3.4	Q60187	methanosarc	215	6	3.4	787	1	YFI3_YEAST	P43596	saccharomyc
143	1	CBSA_SULAC	462	1	3.4	O54088	sulfolobus	216	6	3.4	787	1	Y091_CAEEL	P41842	caenorhabdi
144	1	PPAL_PICPA	468	1	3.4	P52291	plichia past	217	6	3.4	793	1	FURI_RAT	P23377	rattus norv
145	1	GIGA_CHLPN	476	1	3.4	Q926v8	chlamydia p	218	6	3.4	796	1	ABAA_EMENI	P20945	emerlicella
146	1	GABT_CAEEL	483	1	3.4	Q21217	caenorhabdi	219	6	3.4	806	1	SYFB_MYCCE	P47437	mycoplasma
147	1	THIL_PSEAE	484	1	3.4	Q9hu66	pseudomonas	220	6	3.4	811	1	PLSB_VIBCH	Q9kvp8	vibrio chol
148	1	CHET_HUMAN	490	1	3.4	O00409	homo sapien	221	6	3.4	812	1	FAED_ECOLI	P06970	escherichia
149	1	THSA_SULAC	490	1	3.4	Q9v2t5	sulfolobus	222	6	3.4	833	1	CAFA_YERPE	P28949	yersinia pe
150	1	YC46_PORPU	491	1	3.4	P51189	porphyra pu	223	6	3.4	843	1	RPOC_ODOSI	P49467	odonteila s
151	1	YJEM_ECOLI	500	1	3.4	P39282	escherichia	224	6	3.4	861	1	GCR3_YEAST	P34160	saccharomyc
152	1	YC46_GUITH	506	1	3.4	O78439	guillardia	225	6	3.4	873	1	CPHA_SYNY3	P37383	synecocyst
153	1	YML4_YEAST	510	1	3.4	O03750	saccharomyc	226	6	3.4	903	1	DPOL_BPR69	Q38087	bacterioph
154	1	BNA1_HUMAN	512	1	3.4	Q16515	homo sapien	227	6	3.4	906	1	RPOL_BPK11	P18147	bacterioph
155	1	BNA1_RAT	512	1	3.4	Q62962	rattus norv	228	6	3.4	908	1	TB12_NEIMB	Q06987	neisseria m
156	1	VS14_TRYBB	514	1	3.4	P26329	trypanosoma	229	6	3.4	918	1	CAPP_CORGL	P12880	corynebacte
157	1	Y967_TREPA	517	1	3.4	O83933	treponema p	230	6	3.4	927	1	NED4_HUMAN	P46934	homo sapien
158	1	LNT2_RHIME	529	1	3.4	Q52910	rhizobium m	231	6	3.4	936	1	PMP7_CHLPN	Q92898	chlamydia p
159	1	NOP5_HUMAN	529	1	3.4	Q9y2x3	homo sapien	232	6	3.4	941	1	DNAB_RHOMR	Q30477	rhodothermu
160	1	LNT_RHILO	530	1	3.4	Q98b36	rhizobium l	233	6	3.4	953	1	LKAI_PASHA	P16535	pasteurella
161	1	LNT_RHIME	531	1	3.4	P58377	rhizobium m	234	6	3.4	957	1	NED4_MOUSE	P46935	mus musculu
162	1	MLO_HORYU	533	1	3.4	P93766	hordeum vul	235	6	3.4	960	1	GBR1_MOUSE	Q9wv18	mus musculu
163	1	CYDA_AZOVI	537	1	3.4	Q09049	azotobacter	236	6	3.4	961	1	GBR1_HUMAN	Q9ub55	homo sapien
164	1	DEXB_STREQ	537	1	3.4	Q59905	streptococc	237	6	3.4	977	1	A2A1_HUMAN	Q95782	h adapter-r
165	1	EN27_HUMAN	551	1	3.4	Q9bzr9	homo sapien	238	6	3.4	991	1	GBR1_RAT	Q920v4	rattus norv
166	1	EN27_MOUSE	551	1	3.4	Q99p32	mus musculu	239	6	3.4	992	1	POLS_RUBVM	P08563	rubella vir
167	1	TF3B_KLULA	556	1	3.4	P46070	kluyveromyc	240	6	3.4	1014	1	NANH_CLOSE	P29767	clostridium
168	1	G6PI_KLULA	557	1	3.4	P12341	kluyveromyc	241	6	3.4	1018	1	CONT_HAETN	Q12860	homo sapien
169	1	MDLI_PRUSE	563	1	3.4	P52706	prunus sero	242	6	3.4	1028	1	FDXG_HAETN	P46448	haemophilus
170	1	MEND_HAEIN	568	1	3.4	P44612	h menaquin	243	6	3.4	1046	1	HUUA_HAETN	Q48153	haemophilus
171	1	ILVB_LACLA	575	1	3.4	Q02137	lactococcus	244	6	3.4	1054	1	IF2_STIAN	P55875	stigmatella
172	1	SYQ_RALSO	580	1	3.4	Q8y199	raistonia s	245	6	3.4	1063	1	POLS_RUBVI	P21480	rubella vir
173	1	KAPP_ARATH	581	1	3.4	P46014	arabidopsis	246	6	3.4	1077	1	HGPA_HAETN	Q9za21	haemophilus
174	1	YHD9_YEAST	585	1	3.4	P38732	saccharomyc	247	6	3.4	1090	1	PULA_KLEPN	P07206	klebsiella
175	1	LACP_KLULA	587	1	3.4	P07921	kluyveromyc	248	6	3.4	1096	1	PULA_KLEPE	P07811	klebsiella
176	1	PRIM_CLOAB	596	1	3.4	P33655	clostridium	249	6	3.4	1103	1	VG37_BPAR1	Q9q0b5	bacterioph
177	1	TF3B_YEAST	596	1	3.4	P29056	saccharomyc	250	6	3.4	1118	1	YIP1_YEAST	P40456	saccharomyc
178	1	HEXA_YEAST	598	1	3.4	P48823	alteromonas	251	6	3.4	1134	1	KDGS_MESAU	Q40398	mesocricetu
179	1	YMH1_YEAST	607	1	3.4	Q04632	saccharomyc	252	6	3.4	1195	1	CHS4_NEUCR	Q01285	neurospora

253	6	3.4	1216	1	RPOB_TROWH	Q93qf2 tropheryma	326	5	2.8	82	1	PSBE_TOBAC	P05168 nicotiana t
254	6	3.4	1250	1	BXE_CLOBO	Q00496 clostridium	327	5	2.8	82	1	PSBE_WHEAT	P05169 triticum ae
255	6	3.4	1250	1	BXE_CLOBO	P03095 clostridium	328	5	2.8	82	1	YWDA_BACSU	P39609 bacillus su
256	6	3.4	1286	1	IRBP_BOVIN	P12661 bos taurus	329	5	2.8	83	1	HPIIS_CHRGR	P00262 chromatium
257	6	3.4	1290	1	RPOC_MYCPN	P75271 mycoplasma	330	5	2.8	83	1	PSBE_ARATH	P56779 arabidopsis
258	6	3.4	1296	1	BXC_CLOBO	Q60393 clostridium	331	5	2.8	83	1	PSBE_CHLVU	P56309 chlorella v
259	6	3.4	1338	1	ADO_HUMAN	Q06278 homo sapien	332	5	2.8	83	1	PSBE_GUITH	O74846 guillardia
260	6	3.4	1371	1	UBPU_HUMAN	Q94966 homo sapien	333	5	2.8	83	1	PSBE_ODOSI	P49473 odontella s
261	6	3.4	1375	1	GTFC_STRMU	P13470 streptococc	334	5	2.8	83	1	PSBE_PORPU	P51391 porphyra pu
262	6	3.4	1474	1	SHK2_RAT	Q9qx74 rattus norv	335	5	2.8	84	1	PSBE_CVACA	O9tm20 cyanidium c
263	6	3.4	1476	1	CFTR_MOUSE	P26361 mus musculu	336	5	2.8	85	1	FCBG_BOVIN	O9bdr7 bos taurus
264	6	3.4	1476	1	GTFB_STRMU	P08987 streptococc	337	5	2.8	85	1	RR16_TOBAC	O06374 nicotiana t
265	6	3.4	1597	1	SOL_DROME	P27398 drosophila	338	5	2.8	86	1	FCBG_CAVPO	P07249 cavia porce
266	6	3.4	1679	1	YMF9_YEAST	Q04958 saccharomyc	339	5	2.8	86	1	FCBG_HUMAN	P30273 homo sapien
267	6	3.4	2051	1	FAS1_YEAST	P07149 s fatty aci	340	5	2.8	86	1	FCBG_MOUSE	P20491 mus musculu
268	6	3.4	2073	1	FAS1_SCHPO	Q9uu00 s fatty aci	341	5	2.8	86	1	FCBG_PIG	O9xsc6 sus scrofa
269	6	3.4	2133	1	FAB_PIG	P12263 sus scrofa	342	5	2.8	86	1	FCBG_RAT	P20411 rattus norv
270	6	3.4	2210	1	RRPL_EBOSM	Q66802 ebola virus	343	5	2.8	86	1	YABO_BACSU	P37557 bacillus su
271	6	3.4	2319	1	FAB_MOUSE	Q06194 mus musculu	344	5	2.8	86	1	YABO_ECOLI	P21831 escherichia
272	6	3.4	2351	1	FAB_HUMAN	P00451 homo sapien	345	5	2.8	87	1	CEDA_ECOLI	P76211 escherichia
273	6	3.4	3298	1	PC16_HUMAN	Q96190 homo sapien	346	5	2.8	88	1	RR16_SOLUT	P32087 solanum tub
274	6	3.4	3433	1	UTRO_HUMAN	P46939 homo sapien	347	5	2.8	88	1	YPB4_ECOLI	P03853 escherichia
275	6	3.4	4092	1	DYHC_YEAST	P36022 saccharomyc	348	5	2.8	90	1	HS72_USTMA	P18694 ustilago ma
276	6	3.4	4351	1	FAT2_RAT	O88277 rattus norv	349	5	2.8	90	1	Y530_BUCAI	P57556 buchnera ap
277	6	3.4	4639	1	DYHC_DROME	P37276 drosophila	350	5	2.8	90	1	YI44_ARCFU	O28434 archaeoglob
278	5	2.8	20	1	CPA7_PAPSP	P80055 papio sp. (351	5	2.8	91	1	MERP_PSEAE	P04131 pseudomonas
279	5	2.8	22	1	HS71_LEITA	P55938 leishmania	352	5	2.8	91	1	MERP_SHIFL	P04129 shigella fl
280	5	2.8	24	1	PCL1_PACGO	P82421 pachycondyl	353	5	2.8	91	1	RL31_HALNI	O9hmo0 halobacteri
281	5	2.8	24	1	PCL2_PACGO	P82422 pachycondyl	354	5	2.8	91	1	Y1PP_KLULA	P13999 kluyveromyc
282	5	2.8	33	1	YC12_EUGGR	P31559 euglena gra	355	5	2.8	92	1	IHFE_VIBCH	O9kq44 vibrio chol
283	5	2.8	36	1	PSBI_ARATH	P09970 arabidopsis	356	5	2.8	93	1	S10A_BOVIN	P22639 bos taurus
284	5	2.8	36	1	PSBI_HORVU	P25876 hordeum vul	357	5	2.8	93	1	S10A_HUMAN	P20297 homo sapien
285	5	2.8	36	1	PSBI_MARPO	P09369 marchantia	358	5	2.8	93	1	S10A_MOUSE	P56565 mus musculu
286	5	2.8	36	1	PSBI_ORISA	P12161 oryza sativ	359	5	2.8	94	1	GYRA_STAEP	P54112 staphylococ
287	5	2.8	36	1	PSBI_PINTH	P41599 pinus thunb	360	5	2.8	94	1	Y950_TREPA	O83916 treponema p
288	5	2.8	36	1	PSBI_PSEMZ	P29796 pseudotsuga	361	5	2.8	95	1	YCZJ_BACSU	O31484 bacillus su
289	5	2.8	39	1	PSBI_SYNY3	Q55354 synechocyst	362	5	2.8	96	1	MPG3_DAGGL	P93124 dactylis gl
290	5	2.8	44	1	MFW2_SCHPO	P34069 schizosacch	363	5	2.8	97	1	FTRV_MAIZE	P80680 zea mays (m
291	5	2.8	44	1	PHRA_BACSU	Q00829 bacillus su	364	5	2.8	99	1	CX41_MANSP	O46587 mandrillus
292	5	2.8	47	1	YQ54_BACAN	Q9rmx8 bacillus an	365	5	2.8	99	1	REGN_BPH3	O46588 presbytis c
293	5	2.8	49	1	RB12_MOUSE	P35283 mus musculu	366	5	2.8	99	1	REGN_BPH3	P07243 bacterioph
294	5	2.8	49	1	VG12_SPLIR	P15903 spiriplasma	367	5	2.8	99	1	Y162_STRPY	O99y56 streptococc
295	5	2.8	52	1	RK32_GUITH	O78434 guillardia	368	5	2.8	100	1	UCR9_EUGGR	P43266 euglena gra
296	5	2.8	53	1	RS14_METJA	P54110 methanococc	369	5	2.8	101	1	RL36_ENTCP	O91988 enteromorph
297	5	2.8	55	1	CX41_CEBAP	O46589 cebus apell	370	5	2.8	102	1	CCAC_AMPCE	P55738 amphidilum
298	5	2.8	55	1	CX41_SATUS	O46590 saimir ust	371	5	2.8	102	1	RPOL_AERPE	O99y51 aeropyrum p
299	5	2.8	56	1	MT_SYNVU	P30565 synechococc	372	5	2.8	102	1	YK85_CAEEL	P34351 caenorhabdl
300	5	2.8	56	1	RL32_MYCGE	P47603 mycoplasma	373	5	2.8	103	1	NDOA_PSEAE	O51493 pseudomonas
301	5	2.8	59	1	EA57_HUMAN	O43247 homo sapien	374	5	2.8	103	1	NDOA_PSEPU	P23082 pseudomonas
302	5	2.8	64	1	YC33_ODOSI	P49532 odontella s	375	5	2.8	104	1	CYC_RANCA	P00024 rana catesb
303	5	2.8	65	1	YRP2_SYNPY	Q02191 synechococc	376	5	2.8	104	1	RECA_STRSL	P49987 streptococc
304	5	2.8	68	1	RL29_STRPN	O9wv48 streptococc	377	5	2.8	104	1	RL3E_SULAC	P11522 sulfobolus
305	5	2.8	68	1	RL29_STRPY	O9a1w6 streptococc	378	5	2.8	104	1	YBK0_YEAST	P38168 saccharomyc
306	5	2.8	70	1	R37A_SULSO	Q97243 sulfobolus	379	5	2.8	106	1	RLA3_YEAST	P10622 saccharomyc
307	5	2.8	70	1	R37A_SULTO	Q975h0 sulfobolus	380	5	2.8	107	1	COYT_BACSU	P11863 bacillus su
308	5	2.8	71	1	YL32_HALHA	Q06847 halobacteri	381	5	2.8	108	1	YGGL_ECOLI	P38521 escherichia
309	5	2.8	71	1	ARRD_RAT	P36577 rattus norv	382	5	2.8	109	1	YXIH_BACSU	P42300 bacillus su
310	5	2.8	73	1	PSBE_CVACA	P19152 cyanophora	383	5	2.8	110	1	RLA4_YEAST	P02400 saccharomyc
311	5	2.8	75	1	YF45_MYCTU	Q10781 mycobacteri	384	5	2.8	112	1	CFP6_MYCTU	O53251 mycobacteri
312	5	2.8	77	1	HYEP_PSESP	P80048 pseudomonas	385	5	2.8	113	1	CYC_CRIPA	P00078 crithidia f
313	5	2.8	78	1	KEAL_ECOLI	Q52278 escherichia	386	5	2.8	113	1	GVKT_HALNI	P24375 halobacteri
314	5	2.8	78	1	YO09_BPL2	P42544 bacterioph	387	5	2.8	113	1	RL19_MYCLE	O33020 mycobacteri
315	5	2.8	80	1	PSBE_MESVI	Q9muq0 mesostigma	388	5	2.8	113	1	RL19_MYCTU	Q10792 mycobacteri
316	5	2.8	80	1	PSBE_SYNY3	P09190 synechocyst	389	5	2.8	113	1	RLA1_DICDI	P22684 dictyosteli
317	5	2.8	82	1	DNAL_STRAL	P96457 streptomyc	390	5	2.8	113	1	Y186_FOWPV	Q9j548 fowlpox vir
318	5	2.8	82	1	PSBE_CHLRE	P48268 chlamydomon	391	5	2.8	113	1	Y655_BACHD	Q929w1 bacillus ha
319	5	2.8	82	1	PSBE_MARPO	P06851 marchantia	392	5	2.8	114	1	Y655_BACHD	Q929w1 bacillus ha
320	5	2.8	82	1	PSBE_MESCR	P36442 mesembryant	393	5	2.8	114	1	RLA2_EIMTE	Q967y9 elmerella v
321	5	2.8	82	1	PSBE_OENBE	P05170 oenothera b	394	5	2.8	115	1	CMGE_BACSU	P25957 bacillus su
322	5	2.8	82	1	PSBE_OENHO	Q9mtk5 oenothera h	395	5	2.8	115	1	PTHY_RAT	P04089 rattus norv
323	5	2.8	82	1	PSBE_PEA	P13554 pisum sativ	396	5	2.8	115	1	RL19_BACSU	O31742 bacillus su
324	5	2.8	82	1	PSBE_PINTH	P41615 pinus thunb	397	5	2.8	116	1	NU3M_MACRO	P92666 macropus ro
325	5	2.8	82	1	PSBE_SPIOL	P09197 spinacia oi	398	5	2.8	116	1	RL19_MACRO	P30529 bacillus st

837	5	2.8	210	1	HPRT_TRYBB	Q07010	trypanosoma	910	5	2.8	226	1	OAZ_RAT	P54370	rattus norv
838	5	2.8	210	1	PDX5_MOUSE	P99029	mus musculus	911	5	2.8	226	1	VATE_HETSC	Q3513	rhodobacter
839	5	2.8	210	1	UPP_DEIRA	Q9ru32	deinococcus	912	5	2.8	227	1	CBBY_RHOCA	Q3513	rhodobacter
840	5	2.8	211	1	NADD_MYCTU	O86328	mycobacteri	913	5	2.8	227	1	COX2_CHEME	P98027	cheliogaler
841	5	2.8	211	1	NADD_VIBCH	Q9kp38	vibrio chol	914	5	2.8	227	1	COX2_EULMA	P98033	eulemur mac
842	5	2.8	212	1	IL6_CERTO	P46650	cercocebus	915	5	2.8	227	1	COX2_GALSE	P50888	galago sene
843	5	2.8	212	1	IL6_HUMAN	P05231	homo sapien	916	5	2.8	227	1	COX2_HAPGR	P98034	hapalemur g
844	5	2.8	212	1	IL6_MACFA	P79341	macaca fasc	917	5	2.8	227	1	COX2_LEMCA	P98035	lemur catta
845	5	2.8	212	1	IL6_MACMU	P51494	macaca mula	918	5	2.8	227	1	COX2_LEMVA	P98047	lemur varle
846	5	2.8	212	1	IL6_PIG	P26893	sus scrofa	919	5	2.8	227	1	COX2_PROTA	P98042	proplithecus
847	5	2.8	212	1	RAN_DICDI	P33519	dictyostell	920	5	2.8	227	1	RISB_ARATH	O80575	arabidopsis
848	5	2.8	213	1	ACD3_PSEAE	Q9hz17	pseudomonas	921	5	2.8	227	1	YMYO_YEAST	O04304	saccharomyc
849	5	2.8	213	1	INL6_HUMAN	Q9Y581	homo sapien	922	5	2.8	228	1	CLPR_SYNP7	Q9114p4	synchococc
850	5	2.8	213	1	PDX5_RAT	Q9r063	rattus norv	923	5	2.8	228	1	IF6_METJA	P12077	leishmania
851	5	2.8	213	1	PIT_STRHA	P41132	streptomyce	924	5	2.8	228	1	IF6_METJA	P60357	methanococc
852	5	2.8	213	1	RABGR_XENLA	P31412	dictyostell	925	5	2.8	228	1	MODB_RHOCA	O08382	rhodobacter
853	5	2.8	213	1	SKGR_XENLA	P13673	xenopus lae	926	5	2.8	228	1	Y516_BORBU	O51468	borrella bu
854	5	2.8	213	1	YJH2_YEAST	P40359	saccharomyc	927	5	2.8	228	1	Y972_METJA	Q58382	methanococc
855	5	2.8	214	1	CMPA_BACSU	P14204	bacillus su	928	5	2.8	228	1	Y972_METJA	O07215	mycobacter
856	5	2.8	214	1	GSI_HUMAN	Q08623	homo sapien	929	5	2.8	229	1	MTN_HAEIN	P45113	haemophilus
857	5	2.8	214	1	HSTA_CROSI	P02826	drosophila	930	5	2.8	229	1	PRL_MEUGA	P17572	meleagris g
858	5	2.8	214	1	MAPA_CAMJE	Q46122	campylobact	931	5	2.8	229	1	RAS2_NEUCR	Q01387	neurospora
859	5	2.8	214	1	NADD_MYCLE	Q9cbz8	mycobacteri	932	5	2.8	229	1	SFSA_THETN	Q8r7n6	thermoanaer
860	5	2.8	214	1	PDX5_HUMAN	P30044	h peroxidred	933	5	2.8	229	1	UCK_CAEEL	Q17413	caenorhabdl
861	5	2.8	214	1	PCPT_MOUSE	P53808	mus musculus	934	5	2.8	229	1	Y454_TREPA	O83467	treponema p
862	5	2.8	214	1	PRD2_RAT	P53402	arabidopsis	935	5	2.8	229	1	YD74_MYCPN	P75407	mycoplasma
863	5	2.8	215	1	ERD3_ARATH	P35402	arabidopsis	936	5	2.8	230	1	AG11_ARATH	Q38836	arabidopsis
864	5	2.8	215	1	HEX8_ADEMI	P19722	mouse adeno	937	5	2.8	230	1	Y36_YEAST	P53374	saccharomyc
865	5	2.8	215	1	KAD_MYCPN	Q50299	mycoplasma	938	5	2.8	230	1	YLMB_BACSU	O31727	bacillus su
866	5	2.8	215	1	PDX5_CERAE	Q9g1w7	cercoptithec	939	5	2.8	230	1	YSHA_SALTY	Q911r3	salmonella
867	5	2.8	215	1	PDX5_PAPHA	Q9g1w9	papilio hamad	940	5	2.8					

983 5 2.8 239 1 YPCL_ECOLI P18351 escherichia
 984 5 2.8 240 1 HAAL_XENLA Q08821 xenopus lae
 985 5 2.8 240 1 LEC_BOWMI P42088 bowringia m
 986 5 2.8 240 1 SFSA_PASMU P57871 pasteurella
 987 5 2.8 240 1 SODE_HUMAN P08294 homo sapien
 988 5 2.8 240 1 T341_ECOLI P11257 escherichia
 989 5 2.8 240 1 VP26_NPVAC P08358 autographa
 990 5 2.8 241 1 PSA5_HUMAN P28066 homo sapien
 991 5 2.8 241 1 PSA5_MOUSE Q922ul mus musculu
 992 5 2.8 241 1 PSA5_RAT P34064 rattus norv
 993 5 2.8 242 1 6PGL_PSEPU Q9ev79 pseudomonas
 994 5 2.8 242 1 BIOD_BACHD Q9ker5 bacillus ha
 995 5 2.8 243 1 MIP_LEGMI P11106 legionella
 996 5 2.8 243 1 NUKS_HUMAN Q9nie3 homo sapien
 997 5 2.8 243 1 NUKS_RAT Q9epj0 rattus norv
 998 5 2.8 243 1 YKJA_BACSU P49853 bacillus su
 999 5 2.8 243 1 ZIPA_XYLFA Q9pag1 xyella fas
 1000 5 2.8 244 1 ATPQ_IPOBA P22778 ipomoea bat

ALIGNMENTS

RESULT 1

OM12_HAEIN STANDARD; PRT; 459 AA.
 AC P43836;
 DT 01-NOV-1995 (Rel. 32., Created)
 DT 01-NOV-1995 (Rel. 32., Last sequence update)
 DT 15-JUN-2002 (Rel. 41., Last annotation update)
 DE Outer membrane protein P1 precursor (OMP P1).
 GN OMP1 OR HI0401.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 RN [2]
 RP IDENTIFICATION BY MASS SPECTROMETRY.
 RX MEDLINE=20137486; PubMed=10675023;
 RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 RA Gray C., Fountoulakis M.;
 RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
 RL Electrophoresis 21:411-429(2000).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32723; AAC22060.1; -;
 DR TIGR; HI0401; -;
 KW Outer membrane; Signal; Transmembrane; Complete proteome.

FT SIGNAL 1 22 BY SIMILARITY.
 FT CHAIN 23 459 OUTER MEMBRANE PROTEIN P1.
 SQ SEQUENCE 459 AA; 49477 MW; 43BDC98ESA39366F CRC64;
 Query Match 7.38; Score 13; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.le-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 121-GATYKFTPNLSVD 133
 DB 397 GATYKFTPNLSVD 409
 RESULT 2
 OM12_HAEIN STANDARD; PRT; 459 AA.
 ID OM12_HAEIN
 AC P10641;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Outer membrane protein P1 precursor (OMP P1).
 GN OMP1.
 OS Haemophilus Influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=Serotype B;
 RX MEDLINE=88314258; PubMed=2842261;
 RA Munson R.S. Jr., Grass S.;
 RT "Purification, cloning, and sequence of outer membrane protein P1 of
 RT Haemophilus influenzae type b.";
 RL Infect. Immun. 56:2235-2242(1988).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J03381; AAA24990.1; -;
 DR PIR; A28787; A28787.
 DR PIR; A30510; A30510.
 KW Outer membrane; Signal; Transmembrane.
 FT SIGNAL 1 22
 FT CHAIN 23 459 OUTER MEMBRANE PROTEIN P1.
 SQ SEQUENCE 459 AA; 49956 MW; F732CF353CF73F6A CRC64;
 Query Match 7.38; Score 13; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.le-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 121 GATYKFTPNLSVD 133
 DB 397 GATYKFTPNLSVD 409
 RESULT 3
 YCHO_ECOLI STANDARD; PRT; 417 AA.
 ID YCHO_ECOLI
 AC P39165; P46124; P76022;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein ycho.
 GN YCHO OR B1220.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
[3]
RN SEQUENCE OF 1-37 FROM N.A.
RP STRAIN=NM8191;
RC Ivey D.M., Guffanti A.A., Zemsky J., Pinner E., Karpel R.,
RA Padan E., Schuldiner S., Krulwich T.A.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 306-417 FROM N.A.
RP STRAIN=K12;
RX MEDLINE=89263708; PubMed=2657652;
RA Nohno T., Noji S., Taniguchi S., Saito T.;
RT "The narX and narL genes encoding the nitrate-sensing regulators of
Escherichia coli are homologous to a family of prokaryotic two-
component regulatory genes.";
RL Nucleic Acids Res. 17:2947-2957(1989).
[5]
RN SEQUENCE OF 364-417 FROM N.A.
RP STRAIN=K12;
RX MEDLINE=89197802; PubMed=2649492;
RA Stewart V., Parales J. Jr., Merkel S.M.;
RT "Structure of genes narL and narX of the nar (nitrate reductase)
locus in Escherichia coli K-12.";
RL J. Bacteriol. 171:2229-2234(1989).
[6]
RN IDENTIFICATION.
RP MEDLINE=96032851; PubMed=7567469;
RX Borodovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,
RA Danchin A.;
RT "Detection of new genes in a bacterial genome using Markov models for
three gene classes.";
RL Nucleic Acids Res. 23:3554-3562(1995).
CC -1- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.
CC -1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
FRAMESHIFT IN POSITION 398.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000220; AAC74304.1; -;
DR EMBL; D90757; BAA36088.1; -;
DR EMBL; L28709; -; NOT_ANNOTATED_CDS.
DR EMBL; X13360; -; NOT_ANNOTATED_CDS.
DR EMBL; M24910; -; NOT_ANNOTATED_CDS.

DR EMBL; X69189; -; NOT_ANNOTATED_CDS.
DR EcoGene; EGI2405; ycho.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 47023 MW; 4D4841DDC1BE5485 CRC64;

Query Match 4.5%; Score 8; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 GLNLNRYRF 179
Db 256 GLNLNRYRF 263
|||||||

RESULT 4
ID MYG_GOBGI STANDARD; PRT; 146 AA.
AC Q9DEN9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myoglobin.
GN MB.
OS Gobionotothen gibberifrons (Humped rockcod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Gobionotothen.
OX NCBI_TaxID=36202;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart ventricle;
RX MEDLINE=97430285; PubMed=9284559;
RA Vayda M.E., Small D.J., Yuan M.-L., Costello L., Sideli B.D.;
RT "Conservation of the myoglobin gene among Antarctic notothenioid
fishes.";
RL Mol. Mar. Biol. Biotechnol. 6:207-216(1997).
CC -1- FUNCTION: SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES
CC THE MOVEMENT OF OXYGEN WITHIN MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U71057; AAG16645.1; -;
DR HSSP; P02205; LMVT.
DR InterPro; IPR000971; Globin.
DR InterPro; IPR002335; Myoglobin.
DR Pfam; PF00042; globin; 2.
DR PRINTS; PRO0613; MYOGLOBIN.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Muscle.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 59 59 IRON (HEME DISTAL LIGAND).
FT METAL 88 88 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 146 AA; 15576 MW; 6DAAEF3DEB3CE109 CRC64;

Query Match 3.9%; Score 7; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VEADYTT 164
Db 13 VEADYTT 19
|||||||

RESULT 5
MYG_THUAL

ID MYG_THUAL STANDARD; PRT; 146 AA.
AC P02205;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myoglobin.
GN MB.
OS Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8236;
RN [1]
RP SEQUENCE.
RX MEDLINE=81046956; PubMed=7430163;
RA Watts D.A., Rice R.H., Brown W.D.;
RT "The primary structure of myoglobin from yellowfin tuna (Thunnus
RT albacares).";
RL J. Biol. Chem. 255:10916-10924(1980).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.74 ANGSTROMS).
RA Birnbaum G.I., Evans S.V., Przybylska M., Rose D.R.;
RT "1.70-A resolution structure of myoglobin from yellowfin tuna. An
RT example of a myoglobin lacking the D helix.";
RL Acta Crystallogr. D 50:283-289(1994).
CC -1- FUNCTION: SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES
CC THE MOVEMENT OF OXYGEN WITHIN MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A02526; MYTUJ.
DR PB; IMYT; 31-OCT-93.
DR InterPro: IPR000971; Globin.
DR InterPro: IPR002335; Myoglobin.
DR Pfam: PF00042; globin; 1.
DR PRINTS: PR00613; MYOGLOBIN.
DR PROSITE: PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Muscle; Acetylation;
KW 3D-structure.
FT MOD_RES 1 1
FT METAL 59 59 ACETYLATION.
FT 59 59 IRON (HEME DISTAL LIGAND)
FT (BY SIMILARITY).
FT METAL 88 88 IRON (HEME PROXIMAL LIGAND)
FT (BY SIMILARITY).
FT 2 8
FT TURN 9 10
FT HELIX 11 14
FT TURN 15 16
FT HELIX 17 31
FT TURN 33 38
FT TURN 40 44
FT HELIX 47 50
FT TURN 51 52
FT HELIX 54 72
FT TURN 73 75
FT HELIX 78 90
FT TURN 91 91
FT TURN 96 96
FT HELIX 97 113
FT TURN 118 142
FT TURN 143 143
SQ SEQUENCE 146 AA; 13529 MW; 74C72B21D005BD44 CRC64;
Query Match 3.9%; Score 7; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 VEADYTT 164
Db 13 VEADYTT 19
RESULT 6
MYG_THUTH

ID MYG_THUTH STANDARD; PRT; 146 AA.
AC Q9DD47;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myoglobin.
GN MB.
OS Thunnus thynnus (Bluefin tuna), and
OS Thunnus thynnus orientalis (North Pacific bluefin tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8237, 8238;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=T.thynnus, and T.t.orientalis; TISSUE=Skeletal muscle;
RX MEDLINE=21146139; PubMed=11247835;
RA Marcinek D.J., Bonaventura J., Wittenberg J.B., Block B.A.;
RT "Oxygen affinity and amino acid sequence of myoglobins from
RT endothermic and ectothermic fish.";
RL Am. J. Physiol. 280:R1123-R1133(2001).
CC -1- FUNCTION: SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES
CC THE MOVEMENT OF OXYGEN WITHIN MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF291836; AAC02110.1;
DR EMBL: AF291831; AAC02105.1;
DR HSSP: P02205; IMYT.
DR InterPro: IPR000971; Globin.
DR InterPro: IPR002335; Myoglobin.
DR Pfam: PF00042; globin; 2.
DR PRINTS: PR00613; MYOGLOBIN.
DR PROSITE: PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Muscle.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 59 59 IRON (HEME DISTAL LIGAND).
FT METAL 88 88 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 146 AA; 15497 MW; F99735CEC834509A CRC64;
Query Match 3.9%; Score 7; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 VEADYTT 164
Db 13 VEADYTT 19
RESULT 7
CAH2_MOUSE
ID CAH2_MOUSE STANDARD; PRT; 259 AA.
AC P00920;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Carbonic anhydrase II (EC 4.2.1.1) (Carbonate dehydratase II) (CA-II).
GN CA2 OR CAR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84109569; PubMed=6420240;
QY 158 VEADYTT 164
Db 13 VEADYTT 19
RESULT 6
MYG_THUTH

RA Curtis P.J., Withers E., Demuth D., Watt R., Venta P.J., Tashian R.E.;
 RT "The nucleotide sequence and derived amino acid sequence of cDNA
 RL Gene 25:325-332(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86008276; PubMed=2995362;
 RA Venta P.J., Montgomery J.C., Hewett-Emmett D., Wiebauer K.,
 RA Tashian R.E.;
 RT "Structure and exon to protein domain relationships of the mouse
 RT carbonic anhydrase II gene";
 RL J. Biol. Chem. 260:12130-12135(1985).
 RN [3]
 RP SEQUENCE OF 154-177 AND 213-239 FROM N.A.
 RX MEDLINE=83161023; PubMed=6187736;
 RA Curtis P.J.;
 RT "Cloning of mouse carbonic anhydrase mRNA and its induction in mouse
 RT erythroleukemic cells";
 RL J. Biol. Chem. 258:4459-4463(1983).
 RN [4]
 RP SEQUENCE OF 1-76 FROM N.A.
 RX MEDLINE=86077780; PubMed=3000449;
 RA Venta P.J., Montgomery C., Hewett-Emmett D., Tashian R.E.;
 RT "Comparison of the 5' regions of human and mouse carbonic anhydrase
 RT II genes and identification of possible regulatory elements";
 RL Biochim. Biophys. Acta 826:195-201(1985).
 CC -1- FUNCTION: REVERSIBLE HYDRATATION OF CARBON DIOXIDE.
 CC -1- CATALYTIC ACTIVITY: H₂CO(3) = CO(2) + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC-TYPE CARBONIC ANHYDRASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K00811; AAA37356.1;
 DR EMBL; M81022; AAA37357.1;
 DR EMBL; M81016; AAA37357.1; JOINED.
 DR EMBL; M81017; AAA37357.1; JOINED.
 DR EMBL; M81018; AAA37357.1; JOINED.
 DR EMBL; M81019; AAA37357.1; JOINED.
 DR EMBL; M81020; AAA37357.1; JOINED.
 DR EMBL; M81021; AAA37357.1; JOINED.
 DR PIR; A01143; CRMS2.
 DR PIR; A23900; A23900.
 DR PIR; A20539; A20539.
 DR PIR; B23202; B23202.
 DR HSSP; P00918; 1BV3.
 DR MGD; MGI:88269; Car2.
 DR InterPro; IPR001148; Euk_COanh.
 DR Pfam; PF00194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk_COanh; 1.
 DR PROSITE; PS00162; EUK_CO2_ANHYDRASE; 1.
 KW Lysase; Zinc.
 FT INIT_MET 0 0
 FT METAL 93 93 ZINC (CATALYTIC).
 FT METAL 95 95 ZINC (CATALYTIC).
 FT METAL 118 118 ZINC (CATALYTIC).
 FT CONFLICT 38 38 H -> Q (IN REF. 2).
 FT CONFLICT 212 212 E -> D (IN REF. 2).
 SQ SEQUENCE 259 AA; 28960 MW; 5BD81AD512D3A86B CRC64;
 Query Match 3.9%; Score 7; DB 1; Length 259;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 97 YDKAASK 103
 |||||

Db 50 YDKAASK 56
 RESULT 8
 GPDA_STRCO STANDARD; PRT; 336 AA.
 ID GPDA_STRCO
 AC Q92BS0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-
 DE dependent glycerol-3-phosphate dehydrogenase).
 GN GPSA OR SC05559 OR SC7A1.03.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P)(+) =
 CC glycerone phosphate + NAD(P)H.
 CC -1- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
 CC formation.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL034447; CAA22402.1;
 DR InterPro; IPR001652; NAD_Gly3p_dh.
 DR Pfam; PF01210; NAD_Gly3p_dh; 1.
 DR PRINTS; PR00077; GPDHGRGNASE.
 DR ProDom; PD001649; NAD_Gly3p_dh; 1.
 DR PROSITE; PS00957; NAD_G3PDH; 1.
 KW Phospholipid biosynthesis; Oxidoreductase; NAD; Complete proteome.
 SQ SEQUENCE 336 AA; 34688 MW; 4E7BF65F9C640544 CRC64;
 Query Match 3.9%; Score 7; DB 1; Length 336;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 90 TLRAGLA 96
 |||||
 Db 86 TLRAGLA 92
 RESULT 9
 CMGA_BACSU STANDARD; PRT; 356 AA.
 ID CMGA_BACSU
 AC P25933;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

```
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE COMG operon protein 1.
GN COMGA OR COMG1.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008773; PubMed=2507524;
RA Albano M., Bretling R., Dubnau D.A.;
RT "Nucleotide sequence and genetic organization of the Bacillus
RT subtilis comG operon.";
RL J. Bacteriol. 171:5386-5404(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.T., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Fortelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Seklugh J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP FUNCTION.
RX MEDLINE=98083053; PubMed=9422590;
RA Chung Y.S., Dubnau D.A.;
RT "All seven comG open reading frames are required for DNA binding
RT during transformation of competent Bacillus subtilis.";
RN J. Bacteriol. 180:41-45(1998).
CC -1- FUNCTION: REQUIRED FOR UPTAKE OF DNA BY COMPETENT BACILLUS
CC SUBTILIS CELLS.
CC -1- SUBCELLULAR LOCATION: INNER FACE OF THE MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29691; AAA83367.1; -
CC EMBL; D84432; BAA12533.1; -
CC EMBL; Z99116; CAB14404.1; -
CC PIR; B30338; B30338.
CC Subtilist; BG10483; COMGA.
CC InterPro: IPR001482; GSPIL_E.
CC Pfam; PF00437; GSPIL_E; 1.
CC ProDom; PD000739; GSPIL_E; 1.
CC PROSITE; PS00662; T2SP_E; 1.
CC KW Transport; ATP-binding; Complete proteome.
CC FT NP_BIND 144 151 ATP (POTENTIAL).
CC SQ SEQUENCE 356 AA; 40458 MW; C15840FFB734E266 CRC64;
Query Match 3.9%; Score 7; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 GSLTLKL 19
Db 77 GSLTLKL 83
RESULT 10
HN3B_ORYLA STANDARD; PRT; 415 AA.
AC 042097;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hepatocyte nuclear factor 3-beta (HNF-3b) (Me-HNF3B).
GN HNF3B.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Orange-red;
RA Okamoto H.M., Nakayama I., Nagoya H., Araki K.;
RT "Predicted protein structure of medaka FoxA3 and its expression in
RT polster.";
RL Zool. Sci. 18:823-832(2001).
CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LIVER GENES.
CC INTERACTS WITH THE CIS-ACTING REGULATORY REGIONS OF THESE GENES.
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001572; BAA23579.1; -
CC HSP; Q63245; 2HFH.
CC InterPro: IPR001766; TF_Fork_head.
CC Pfam; PF00250; Fork_head; 1.
CC PRINTS; PR00053; FORKHEAD.
CC ProDom; PD000425; TF_Fork_head; 1.
CC SMART; SM00339; FH; 1.
CC PROSITE; PS00657; FORK_HEAD_1; 1.
CC PROSITE; PS00658; FORK_HEAD_2; 1.
CC PROSITE; PS50039; FORK_HEAD_3; 1.
```


KW DNA-binding: Nuclear protein; Transcription regulation; Activator.
 FT DOMAIN 32 124 MET-RICH.
 FT DNA_BIND 150 244 FORK-HEAD.
 SQ SEQUENCE 415 AA; 45722 MW; CB85A0C2A20FF52A CRC64;

Query Match 3.9%; Score 7; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 SASIPDT 113
 |||||
 Db 394 SASIPDT 400

RESULT 11

OM47_PASMU STANDARD; PRT; 428 AA.
 AC P80603; Q9CLYO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 47 kDa outer membrane protein precursor.
 GN PM1069.

OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=Pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

RN [2]

RP SEQUENCE OF 26-44.

RC STRAIN=A225;

RA Hartmann L.;

RL Submitted (MAY-1996) to the SWISS-PROT data bank.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.

CC -1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE006147; AA03153.1; ALT_INIT.

FW Outer membrane; Signal; Transmembrane; Complete proteome.

KW SIGNAL 1 25

FT CHAIN 26 428 47 KDa OUTER MEMBRANE PROTEIN.

SQ SEQUENCE 428 AA; 45061 MW; D9606AC50F8220EC CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 HYSKYKT 44
 |||||
 Db 298 HYSKYKT 304

RESULT 12

FADL_ECOLI STANDARD; PRT; 448 AA.

AC P10384; P77697;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Long-chain fatty acid transport protein precursor (Outer membrane

DE FADL protein) (Outer membrane flp protein).
 GN FADL OR TTR OR B2344.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-38.
 RX MEDLINE=91100327; PubMed=1987139;
 RA Black P.N.;
 RT "Primary sequence of the Escherichia coli fadL gene encoding an outer
 RT membrane protein required for long-chain fatty acid transport.";
 RL J. Bacteriol. 173:435-442(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97349980; PubMed=9205837;

RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,

RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,

RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,

RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Silvasundaram S.,

RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,

RA Yamagata S., Horiuchi T.;

RT "Construction of a contiguous 874-kb sequence of the Escherichia coli

RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and

RT analysis of its sequence features.";

RL DNA Res. 4:91-113(1997).

RN [4]

RP SEQUENCE OF 66-448 FROM N.A.

RC STRAIN=K12;

RX MEDLINE=88288050; PubMed=2840553;

RA Said B., Ghosh C.R., Vu L., Nunn W.;

RT "Nucleotide sequencing and expression of the fadL gene involved in

RT long-chain fatty acid transport in Escherichia coli.";

RL Mol. Microbiol. 2:363-370(1988).

CC -1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS

CC ACROSS THE OUTER MEMBRANE. IT IS A RECEPTOR FOR THE BACTERIOPHAGE

CC T2. FADL MAY FORM A SPECIFIC CHANNEL.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.

CC CONTAINS SEVERAL POTENTIAL TRANSMEMBRANE DOMAINS.

CC -1- INDUCTION: BY LONG-CHAIN FATTY ACIDS. EXPRESSION OF FADL IS UNDER

CC THE CONTROL OF THE FADR REPRESSOR.

CC -1- SIMILARITY: BELONGS TO THE OMP1/FADL FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M60607; AA64433.1; -

DR EMBL; AE000323; AAC75404.1; -

DR EMBL; D90865; BAA16205.1; -

DR EMBL; Y00552; CAA68630.1; -

DR PIR; A39126; A39126.

DR PIR; S02829; S02829.

DR EcoGene; EG10280; fadL.

KW Outer membrane; Lipid transport; Phage recognition; Signal;

KW Transmembrane; Complete proteome.

FT SIGNAL 1 27

CHAIN 28 448

LONG-CHAIN FATTY ACID TRANSPORT PROTEIN.

```

FT CONFLICT 178 178 A -> V (IN REF. 1).
FT CONFLICT 180 ARA -> RRP (IN REF. 4).
FT CONFLICT 288 R -> A (IN REF. 2 AND 3).
SQ SEQUENCE 448 AA; 48857 MW; BIA48EA07210FA88 CRC64;

Query Match 3.98; Score 7; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 QWAIHYS 40
Db 313 QWAIHYS 319
|||||

RESULT 13
AMID_PSEPÚ STANDARD; PRT; 466 AA.
AC 069768;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Amidase (EC 3.5.1.4).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 18668;
RX MEDLINE=99025625; PubMed=9809753;
RA Wu S., Fallon R.D., Payne M.S.;
RT "Cloning and nucleotide sequence of amidase gene from Pseudomonas
RT putida.";
RL DNA Cell Biol. 17:915-920(1998).
CC -1- CATALYTIC ACTIVITY: A monocarboxylic acid amide + H(2)O -> a
CC monocarboxylate + NH(3).
CC -1- PATHWAY: INVOLVED IN THE METABOLISM OF NITRILES.
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U89363; AAC18422.1; -.
DR InterPro; IPR000120; Amidase.
DR Pfam; PF01425; Amidase; 1.
DR PROSITE; PS00571; AMIDASES; 1.
KW Hydrolyase.
SQ SEQUENCE 466 AA; 50162 MW; AB9FF610B7CF4CF5 CRC64;

Query Match 3.98; Score 7; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LRLAGLAY 97
Db 443 LRLAGLAY 449
|||||

RESULT 14
CD19_MOUSE STANDARD; PRT; 547 AA.
AC P25918;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen CD19 precursor (B-lymphocyte surface antigen
DE B4) (Leu-12) (Differentiation antigen CD19).
CN CD19.

```

```

OS Mus musculus. (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=9132460; PubMed=1714482;
RA Zhou L.J., Ord D.C., Hughes A.L., Tedder T.F.;
RT "Structure and domain organization of the CD19 antigen of human,
RT mouse, and guinea pig B lymphocytes. Conservation of the extensive
RT cytoplasmic domain.";
RL J. Immunol. 147:1424-1432(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=92137946; PubMed=1370948;
RA Zhou L.J., Ord D.C., Omori S.A., Tedder T.F.;
RT "Structure of the genes encoding the CD19 antigen of human and mouse
RT B lymphocytes.";
RL Immunogenetics 35:102-111(1992).
RN [3]
RP SEQUENCE OF 302-547 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=89292699; PubMed=2472450;
RA Tedder T.F., Isaacs C.M.;
RT "Isolation of cDNAs encoding the CD19 antigen of human and mouse B
RT lymphocytes. A new member of the immunoglobulin superfamily.";
RL J. Immunol. 143:712-717(1989).
RN [4]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=92269839; PubMed=1375324;
RA Kozmik Z., Wang S., Doerfler P., Adams B., Busslinger M.;
RT "The promoter of the CD19 gene is a target for the B-cell-specific
RT transcription factor BSAP.";
RL Mol. Cell. Biol. 12:2662-2672(1992).
CC -1- FUNCTION: MAY BE INVOLVED IN GROWTH REGULATION OF B-CELLS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62542; AAA37388.1; -.
DR EMBL; M62553; AAA37390.1; -.
DR EMBL; M62551; AAA37390.1; JOINED.
DR EMBL; M62552; AAA37390.1; JOINED.
DR EMBL; M28240; AAA74753.1; -.
DR EMBL; M84372; AAA37389.1; -.
DR PIR; B45808; B45808.
DR MGD; MGI:88319; Cd19.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 2.
DR Immunoglobulin domain; Glycoprotein; Transmembrane; B-cell; Signal;
KW Repeat.
RN [5]
RP SIGNAL.
FT PROPEP 17 16
FT CHAIN 19 547
FT DOMAIN 19 287
FT TRANSMEM 288 311
FT DOMAIN 312 547
FT DOMAIN 31 104
FT DOMAIN 190 266
FT DISULFID 38 97
FT DISULFID 197 259

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M28170; AAC68490.1; -
DR EMBL: M21097; AAA35533.1; ALT_FRAME.
DR EMBL: M84371; AAA69966.1; -
DR EMBL: M62530; AAB60697.1; -
DR EMBL: M62544; AAB60697.1; JOINED.
DR EMBL: M62545; AAB60697.1; JOINED.
DR EMBL: M62546; AAB60697.1; JOINED.
DR EMBL: M62547; AAB60697.1; JOINED.
DR EMBL: M62548; AAB60697.1; JOINED.
DR EMBL: M62549; AAB60697.1; JOINED.
DR PIR: JLO074; JLO074.
DR PIR: A4441; A4441.
DR Genew: HGNC:1633; CD19.
DR MIM: 107265; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; ig; 2.
DR SMART: SM00409; IG; 2.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; B-cell; Signal;
KW Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 556 B-LYMPHOCYTE ANTIGEN CD19.
FT DOMAIN 20 291 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 292 313 POTENTIAL.
FT DOMAIN 314 556 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 104 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 193 268 IG-LIKE C2-TYPE DOMAIN 2.
FT DISULFID 38 97 POTENTIAL.
FT CARBOHYD 200 261 POTENTIAL.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 29 29 E -> EG (IN REF. 4).
FT CONFLICT 80 80 S -> I (IN REF. 2).
FT CONFLICT 186 186 Q -> QALVLSLPVP (IN REF. 3).
FT CONFLICT 514 514 H -> R (IN REF. 3 AND 4).
FT CONFLICT 514 514 H -> R (IN REF. 3 AND 4).
SQ SEQUENCE 556 AA; 61069 MW; 5957952A169AFB42 CRC64;

Query Match 3.9%; Score 7; DB 1; Length 556; +
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 QDGSYGE 63
DB 427 QDGSYGE 433
|||||||

RESULT 15
CD19_HUMAN
AC P15391; STANDARD; PRT; 556 AA.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen CD19 precursor (B-lymphocyte surface antigen
DE B4) (Leu-12) (Differentiation antigen CD19).
GN CD19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil.
RX MEDLINE=89292699; PubMed=2472450;
RA Tedder T.F., Isaacs C.M.;
RT "Isolation of cDNAs encoding the CD19 antigen of human and mouse B
RT lymphocytes. A new member of the immunoglobulin superfamily.";
RL J. Immunol. 143:712-717(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89010509; PubMed=2459292;
RA Stamenkovic I., Seed B.;
RT "CD19, the earliest differentiation antigen of the B cell lineage,
RT bears three extracellular immunoglobulin-like domains and an
RT Epstein-Barr virus-related cytoplasmic tail.";
RL J. Exp. Med. 168:1205-1210(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92269839; PubMed=1375324;
RA Kozmik Z., Wang S., Doerfler P., Adams B., Busslinger M.;
RT "The promoter of the CD19 gene is a target for the B-cell-specific
RT transcription factor BSAP.";
RL Mol. Cell. Biol. 12:2662-2672(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92137946; PubMed=1370948;
RA Zhou L.J., Ord D.C., Omori S.A., Tedder T.F.;
RT "Structure of the genes encoding the CD19 antigen of human and mouse
RT B lymphocytes.";
RL Immunogenetics 35:102-111(1992).
CC -!- FUNCTION: MAY BE INVOLVED IN GROWTH REGULATION OF B-CELLS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 396
CC ONWARD AND IS SHORTER (467 AA) DUE TO A FRAMESHIFT.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD19 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd19.htm".
CC -----
```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M28170; AAA68490.1; -
DR EMBL: M21097; AAA35533.1; ALT_FRAME.
DR EMBL: M84371; AAA69966.1; -
DR EMBL: M62530; AAB60697.1; -
DR EMBL: M62544; AAB60697.1; JOINED.
DR EMBL: M62545; AAB60697.1; JOINED.
DR EMBL: M62546; AAB60697.1; JOINED.
DR EMBL: M62547; AAB60697.1; JOINED.
DR EMBL: M62548; AAB60697.1; JOINED.
DR EMBL: M62549; AAB60697.1; JOINED.
DR PIR: JLO074; JLO074.
DR PIR: A4441; A4441.
DR Genew: HGNC:1633; CD19.
DR MIM: 107265; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; ig; 2.
DR SMART: SM00409; IG; 2.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; B-cell; Signal;
KW Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 556 B-LYMPHOCYTE ANTIGEN CD19.
FT DOMAIN 20 291 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 292 313 POTENTIAL.
FT DOMAIN 314 556 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 104 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 193 268 IG-LIKE C2-TYPE DOMAIN 2.
FT DISULFID 38 97 POTENTIAL.
FT CARBOHYD 200 261 POTENTIAL.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 29 29 E -> EG (IN REF. 4).
FT CONFLICT 80 80 S -> I (IN REF. 2).
FT CONFLICT 186 186 Q -> QALVLSLPVP (IN REF. 3).
FT CONFLICT 514 514 H -> R (IN REF. 3 AND 4).
FT CONFLICT 514 514 H -> R (IN REF. 3 AND 4).
SQ SEQUENCE 556 AA; 61069 MW; 5957952A169AFB42 CRC64;

Query Match 3.9%; Score 7; DB 1; Length 556; +
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 QDGSYGE 63
DB 434 QDGSYGE 440
|||||||

RESULT 16
FIBP_ADE31
ID FIBP_ADE31 STANDARD; PRT; 556 AA.
AC P36848;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Fiber protein.
GN FIB.
OS Human adenovirus type 31.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10529;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VRL 15/62;
```

```

RX MEDLINE-96151051; PubMed=8578008;
RA Pring-Akerblom P., Adrian T.;
RT "Sequence characterization of the adenovirus 31 fibre and comparison
RL with serotypes of subgenera A to F.";
RS. Virol. 146:343-354(1995).
CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND
CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
CC -1- SUBUNIT: HOMOTRIMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X76548; CAA54050.1; -
CC PR: S39297; S39297.
CC HSP; P11818; IKNB.
CC InterPro: IPR000939; Adeno_fiber2.
CC InterPro: IPR000978; Adeno_fiber_knob.
CC InterPro: IPR000931; Adeno_fiber.
CC Pfam: PF00541; adeno_fiber; 1.
CC Pfam: PF00608; adeno_fiber2; 6.
CC PRINTS: PR00307; ADENOVSFIBRE.
CC Fiber protein.
CC KW FIBER PROTEIN.
CC SQ SEQUENCE 556 AA; 58944 MW; 6FB13A60ED1D0C33 CRC64;
Query Match 3.9%; Score 7; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 GSLPLKL 19
Db 66 GSLPLKL 72
RESULT 17
RPA2 METJA STANDARD; PRT; 859 AA.
AC Q58446;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase subunit A" (EC 2.7.7.6) [Contains: Mja
DE rpoA2 intein (Mja rpol A" intein)].
DE RPOA2 OR MJ1043
GN Methanococcus jannaschii.
OS Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES

```

```

CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67547; AAB99047.1; -
CC TIGR: MJ1043;
CC InterPro: IPR003586; Hedgehog_hintC.
CC InterPro: IPR003587; Hedgehog_hintN.
CC InterPro: IPR002203; Intein.
CC InterPro: IPR004042; Intein_endonuc.
CC Pfam: PF01854; RNA_pol_A2; 2.
CC SMART: SM00305; HintC; 1.
CC SMART: SM00306; HintN; 1.
CC PROSITE: PS50818; INTEIN_CTER; 1.
CC PROSITE: PS50819; INTEIN_NTER; 1.
CC PROSITE: PS50817; INTEIN_NTER; 1.
CC Transcription; Transcription; DNA-directed RNA polymerase; Zinc;
CC Autocatalytic cleavage; Protein splicing; Complete proteome;
CC Hydrolase; Nuclease; 75
FT CHAIN 1 76 546
FT CHAIN 1 MJA RPOA2 INTEIN (SPACER PROTEIN)
FT CHAIN 1 (POTENTIAL).
FT CHAIN 547 859
FT CHAIN RPOA2, 2ND PART (POTENTIAL).
SQ SEQUENCE 859 AA; 97086 MW; A05799E007899015 CRC64;
Query Match 3.9%; Score 7; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 KEEYKDN 74
Db 588 KEEYKDN 594
RESULT 18
HTRE_ECOLI STANDARD; PRT; 865 AA.
AC P33129;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane usher protein htrE precursor (Heat shock protein E).
DE HTRE OR B0139.
GN Escherichia coli.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN-K12;
RX MEDLINE-93332405; PubMed=8102362;
RA Raina S., Missiakas D., Baird L., Kumar S., Georgopoulos C.;
RT "Identification and transcriptional analysis of the Escherichia coli
RT htrE operon which is homologous to pap and related pilin operons.";
RL J. Bacteriol. 175:5009-5021(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE-94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 2.4-4.1 min (110,917-193,643 bp) region.";

```

Nucleic Acids Res. 22:1637-1639(1994).

[3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).

[4]
 RN SEQUENCE OF 30-33.
 RP Raina S.;
 RL Submitted (DEC-1994) to the SWISS-PROT data bank.
 CC -1- FUNCTION: PROBABLE PORIN-LIKE PROTEIN NECESSARY FOR THE ASSEMBLY
 CC OF A PILIN-TYPE PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
 CC (by similarity).
 CC -1- INDUCTION: BY HEAT SHOCK.
 CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 481 WHICH CREATES TWO SEPARATE ORFS.
 CC -1- CAUTION: REF.1 SEQUENCE WAS INCORRECT IN POSITION 861 ONWARD DUE
 CC TO CLONING ARTIFACT.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: L00680; AAA23721.1; ALT-SEQ.
 DR EMBL: D26582; ; NOT-ANNOTATED_CDS.
 DR EMBL: AE000123; AAC73250.1; ;
 DR EcoGene; EGI1972; htrE.
 DR InterPro; IPR000015; Fimb_usher.
 DR Pfam; PF00577; Usher; 1.
 DR PROSITE; PS01151; FIMBRIAL_USHER; 1.
 KW Outer membrane; Transmembrane; Fimbria; Transport; Signal; Heat shock;
 KW Complete proteome. 29
 FT SIGNAL 1
 FT CHAIN 30 865 OUTER MEMBRANE USHER PROTEIN HTRE.
 FT DISULFID 838 862 POTENTIAL.
 FT CONFLICT 51 51 S -> T (IN REF. 1).
 FT CONFLICT 810 811 OG -> HR (IN REF. 1).
 FT CONFLICT 849 849 E -> P (IN REF. 1).
 SQ SEQUENCE 865 AA; 95499 MW; F1748B551E4A5AAE CRC64;

Query Match 3.9%; Score 7; DB 1; Length 865;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 LNDALTL 91
 Db 474 LNDALTL 480
 ID A2AL_MOUSE STANDARD; PRT; 977 AA.
 AC PI7426;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adapter-related protein complex 2 alpha 1 subunit (Alpha-adaptin A)
 DE (Adaptor protein complex Ap-2 alpha-1 subunit) (Clathrin assembly
 DE protein complex 2 alpha-A large chain) (100 kDa coated vesicle protein
 DE A) (Plasma membrane adaptor HA2/AP2 adaptin alpha A subunit).
 GN AP2A1 OR ADTAA OR CLAP1.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-89155572; PubMed-2564002;
 RA Robinson M.S.;
 RT "Cloning of cDNAs encoding two related 100-kD coated vesicle proteins
 RL (alpha-adaptins).";
 RL J. Cell Biol. 108:833-842(1989).

[2]
 RN CHARACTERIZATION OF ISOFORMS A AND B.
 RP PubMed=7593326;
 RX Ball C.L., Hunt S.P., Robinson M.S.;
 RT "Expression and localization of alpha-adaptin isoforms.";
 RL J. Cell Sci. 108:2865-2875(1995).

CC -1- FUNCTION: Subunit of clathrin-associated adaptor protein complex 2
 CC that plays a role in protein sorting in the late-Golgi/trans-Golgi
 CC network (TGN) and/or endosomes. The AP complexes mediate both the
 CC recruitment of clathrin to membranes and the recognition of
 CC sorting signals within the cytosolic tails of transmembrane cargo
 CC molecules. AP-2 complex seems to play a role in the recycling of
 CC synaptic vesicle membranes from the presynaptic surface (by
 CC similarity).

CC -1- SUBUNIT: Adaptor protein complex 2 (AP-2) is a heterotetramer
 CC composed of two large adaptins (alpha1A/AP2A1 or alpha1B/AP2A1 or
 CC alpha2/AP2A2 and beta1/AP2B1), a medium adaptin (mu2/AP2M1) and a
 CC small adaptin (sigma2/AP2S1).

CC -1- SUBCELLULAR LOCATION: Component of the coat surrounding the
 CC cytoplasmic face of coated vesicles located at the Golgi complex
 CC (by similarity).

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
 CC produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: Isoform A is expressed only in neuronal tissue
 CC and skeletal muscle. Isoform B is widely expressed.

CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES LARGE SUBUNITS
 CC FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: X14971; CAA33096.1; ;
 DR PIR; A30111; A30111.
 DR MGD; MGI:101921; Ap2a1.
 DR InterPro; IPR002553; Adaptin_N.
 DR InterPro; IPR003164; Alpha_adaptin_C.
 DR InterPro; IPR001121; Gamma_adaptin_C.
 DR Pfam; PF01602; Adaptin_N; 1.
 DR Pfam; PF02296; Alpha_adaptin_C; 1.
 DR Pfam; PF02883; Alpha_adaptin_C2; 1.
 KW Golgi stack; Protein transport; Transport; Coated pits; Endocytosis;
 KW Alternative splicing.
 FT VARSPIC 706 727 MISSING (IN ISOFORM B).
 SQ SEQUENCE 977 AA; 107663 MW; F4ED87D3F9EF230A CRC64;

Query Match 3.9%; Score 7; DB 1; Length 977;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SLTLKLP 20
 Db 852 SLTLKLP 858
 ID YC52_XYLFA STANDARD; PRT; 1641 AA.
 RESULT 20
 YC52_XYLFA

```

AC Q9PDX7;
DT 15-JUN-2002 (Rel. 41, Created)
DE 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical protein XF1252 precursor..
GN XF1252.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuranai E.E., Laligret F., Lambais M.R., Leite L.C.C.,
RA Lenos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meldanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003959; AAF84061.1;
CC KW Hypothetical protein; signal; Complete proteome.
CC FT SIGNAL 1 31 POTENTIAL.
CC FT CHAIN 32 1641 HYPOTHETICAL PROTEIN XF1252.
CC SQ SEQUENCE 1641 AA; 178920 MW; DF20B9B107364DB3 CRC64;

Query Match 3.9%; Score 7; DB 1; Length 1641;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 KGLLVE 159
Db 773 KGLLVE 779
|||||

RESULT 21
ZN32_HUMAN STANDARD; PRT; 52 AA.
AC P17041;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
Zinc finger protein 32 (Zinc finger protein K0X30) (Fragment).
ZNF32 OR K0X30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RT cells.";
RL New Biol. 2:363-374(1990).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52361; CAA36587.1;
CC PIR; S10420; S10420.
CC DR HSPP; P15822; IBB0.
CC DR Genew; HGNC:13095; ZNF32.
CC MIM; 194539;
CC DR InterPro: IPR000822; Znf_C2H2.
CC Pfam; PF000096; zf_C2H2; 2.
CC ProDom; PD000003; Znf_C2H2; 2.
CC SMART; SM00355; Znf_C2H2; 2.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
CC Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
CC Nuclear protein; Repeat.
CC FT NON_TER 1 1
CC FT ZN_FING 1 23 C2H2-TYPE.
CC FT ZN_FING 29 51 C2H2-TYPE.
CC FT NON_TER 52 52
CC SQ SEQUENCE 52 AA; 6029 MW; 97E552B7C3918F78 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KGSLLL 17
Db 13 KGSLLL 18
|||||

RESULT 22
RS20_LACLA STANDARD; PRT; 77 AA.
AC Q9CEU5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S20.
GN RPST OR LL1739.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,

```

RA Weissbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
 RL *lactis* ssp. *lactis* IL1403.";
 CC Genome Res. 11:731-753(2001).
 CC -!- FUNCTION: Binds directly to 16S ribosomal RNA (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE006404; AAK05837.1; -
 DR InterPro: IPR002583; Ribosomal_S20p.
 DR Pfam: PF01649; Ribosomal_S20p; 1.
 DR ProDom: PD004231; Ribosomal_S20p; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SQ SEQUENCE 77 AA; 8353 MW; 975F5D44BF145174 CRC64;

 Query Match 3.4%; Score 6; DB 1; Length 77;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 98 DRAASK 103
 Db 52 DRAASK 57

 RESULT 23
 RS20_MYCLE
 ID RS20_MYCLE STANDARD; PRT; 86 AA.
 AC Q3132;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S20.
 GN RPST OR ML0604 OR MCL536.06.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Elgmeler K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -!- FUNCTION: Binds directly to 16S ribosomal RNA (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 299125; CAB16149.1; -
 DR EMBL: AL583919; CAC30112.1; -
 DR Lepronia; ML0604; -

DR InterPro: IPR002583; Ribosomal_S20p.
 DR Pfam: PF01649; Ribosomal_S20p; 1.
 DR ProDom: PD004231; Ribosomal_S20p; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SQ SEQUENCE 86 AA; 9609 MW; 13B18F2730582790 CRC64;

 Query Match 3.4%; Score 6; DB 1; Length 86;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 98 DRAASK 103
 Db 59 DRAASK 64

 RESULT 24
 RS20_MYCTU
 ID RS20_MYCTU STANDARD; PRT; 86 AA.
 AC P71731;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S20.
 GN RPST OR RV2412 OR MT2485 OR MTCY253.08C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeler K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleisemann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Woldman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Binds directly to 16S ribosomal RNA (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 281368; CAB03746.1; -
 DR EMBL: AE007087; AAK46781.1; -
 DR TIGR: MT2485; -
 DR Tuberculist; RV2412; -
 DR InterPro: IPR002583; Ribosomal_S20p.
 DR Pfam: PF01649; Ribosomal_S20p; 1.
 DR ProDom: PD004231; Ribosomal_S20p; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.

```
SO SEQUENCE 86 AA; 9405 MW; 6B78BBFABC6DF1AC CRC64;
Query Match 3.4%; Score 6; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 DKAASK 103
| | | | |
Db 59 DKAASK 64

RESULT 25
RL35_PIG STANDARD; PRT; 94 AA.
AC Q29361;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L35 (fragment).
GN RPL35.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
R MEDLINE=96327607; PubMed=8677129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RL Mamm. Genome 7:509-517(1996).
CC -1- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: F14766; CAA23237.1;
DR InterPro: IPR001854; Ribosomal_L29.
DR Pfam: PF00831; Ribosomal_L29; 1.
DR TIGRFAMs: TIGR00012; L29; 1.
DR PROSITE: PS00579; RIBOSOMAL_L29; 1.
KW Ribosomal protein.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10926 MW; 19D8FEBD51F41D3F CRC64;

Query Match 3.4%; Score 6; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRGKKK 144
| | | | |
Db 7 LRGKKK 12

RESULT 26
Y004_BPHPI STANDARD; PRT; 97 AA.
AC P51703;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 8.9 kDa protein in INT-C1 intergenic region (ORF4)
DE (ORF18).
OS Bacteriophage HPI.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC unclassified Myoviridae.

OX NCBI_TaxID=10690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP1C1;
R MEDLINE=95089704; PubMed=7997180;
RA Esposito D., Scocca J.J.;
RT "Identification of an HPI phage protein required for site-specific
RL excision.";
RL Mol. Microbiol. 13:685-695(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HP1C1;
R MEDLINE=96279738; PubMed=8710508;
RA Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA Waldman A.S., Scocca J.J.;
RT "The complete nucleotide sequence of bacteriophage HPI DNA.";
RL Nucleic Acids Res. 24:2360-2368(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U24159; AAB09186.1;
DR Hypothetical protein.
KW SEQUENCE 97 AA; 11266 MW; 7B76B495CEBDC139 CRC64;
SO SEQUENCE 97 AA; 11266 MW; 7B76B495CEBDC139 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LTLKLP 20
| | | | |
Db 56 LTLKLP 61

RESULT 27
RK21_PORPU STANDARD; PRT; 104 AA.
AC P51209;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chloroplast 50S ribosomal protein L21.
GN RPL21.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVONPORT;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RL genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: THIS PROTEIN BINDS TO 23S RIBOSOMAL RNA IN THE PRESENCE
CC OF PROTEIN L20 (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38804; AAC08095.1;
DR InterPro: IPR001787; Ribosomal_L21p.
```



```

DR Pfam; PF00829; Ribosomal_L21p; 1.
DR ProDom; PD003604; Ribosomal_L21p; 1.
DR TIGRFAMS; TIGR00061; L21; 1.
DR PROSITE; PS01169; RIBOSOMAL_L21; 1.
KW Ribosomal protein; rRNA-binding; Chloroplast.
SQ SEQUENCE 104 AA; 11877 MW; 76BE1C85F0DC355 CRC64;

Query Match      3.4%; Score 6; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 HLRGKK 143
DB 67 HLRGKK 72

RESULT 28
ID YHFI_ECOLI          STANDARD;          PRT; 120 AA.
AC P45551;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yhfY.
GN YHFY OR B3382.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: TO H. INFLUENZAE HI0487.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18937; AAC58179.1; ALT_INIT.
CC EMBL; AB000414; AAC76407.1; ALT_INIT.
CC EcoGene; EG12920; yhfY.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 120 AA; 13528 MW; A3FFB6494C17202C CRC64;

Query Match      3.4%; Score 6; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ANLYGL 173
DB 107 ANLYGL 112

RESULT 29
ID RL35_HUMAN          STANDARD;          PRT; 122 AA.
AC P42766;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L35.
GN RPL35.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90211326; PubMed=232279;
RA Suzuki K., Olivera J., Wool I.G.;
RT "The primary structure of rat ribosomal protein L35.";
RL Biochem. Biophys. Res. Commun. 167:1377-1382(1990).
CC -!- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Patel S.K., Chandraratna R., Nagpal S.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12465; AA51648.1; -.
CC EMBL; BC000348; AAH00348.1; -.
CC SWISS-2DPAGE; P42766; HUMAN.
CC Genew; HGNC:10344; RPL35.
CC InterPro; IPR001854; Ribosomal_L29.
CC Pfam; PF00831; Ribosomal_L29; 1.
CC TIGRFAMS; TIGR00012; L29; 1.
CC PROSITE; PS00579; RIBOSOMAL_L29; 1.
KW RIBOSOMAL protein.
FT INIT MET 0 0 BY SIMILARITY.
SQ SEQUENCE 122 AA; 14420 MW; F44742BADAAC725D CRC64;

Query Match      3.4%; Score 6; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRGKKK 144
DB 8 LRGKKK 13

RESULT 30
ID RL35_RAT            STANDARD;          PRT; 122 AA.
AC P17078;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE 60S ribosomal protein L35.
GN RPL35.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90211326; PubMed=232279;
RA Suzuki K., Olivera J., Wool I.G.;
RT "The primary structure of rat ribosomal protein L35.";
RL Biochem. Biophys. Res. Commun. 167:1377-1382(1990).
CC -!- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: X51705; CAA36001.1; -.
DR PIR: S10390; R5RT35.
DR InterPro: IPR001854; Ribosomal_L29.
DR Pfam: PF00831; Ribosomal_L29; 1.
DR TIGRfams: TIGR00012; L29; 1.
DR PROSITE: PS00579; RIBOSOMAL_L29; 1.
KW Ribosomal protein.
FT INIT_MET 0
SQ SEQUENCE 122 AA; 14421 MW; 014742A1B1674EB CRC64;

Query Match 3.4%; Score 6; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LRGGKK 144
DB 8 LRGGKK 13
|||||

RESULT 31
MSP1_GLO
ID MSP1_GLO STANDARD; PRT; 125 AA.
AC P53021;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major sperm protein 1.
GN MSP-1.
OS Globodera rostochiensis (Golden nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX NCBI_TaxID=31243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ro1 / Mierenbos;
RA Novitski C.E., Brown S., Chen R., Corner A.S., Atkinson H.J.,
RA McPherson M.J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC Sperm crawling. Forms an extensive filament system that extends
CC from sperm villipoda, along the leading edge of the pseudopod (by
CC similarity).
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24499; AAA29146.1; -.
CC HSSP: P27439; 1MSP.
CC InterPro: IPR000535; MSP_domain.
CC Pfam: PF00635; MSP_domain; 1.
CC Cytokeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13858 MW; 0255C31F187549BC CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8
DB 58 NGVLGP 63
|||||

RESULT 32
MSP3_GLO
ID MSP3_GLO STANDARD; PRT; 125 AA.
AC P53023;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major sperm protein 3.
GN MSP-3.
OS Globodera rostochiensis (Golden nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX NCBI_TaxID=31243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ro1 / Mierenbos;
RA Novitski C.E., Brown S., Chen R., Corner A.S., Atkinson H.J.,
RA McPherson M.J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC Sperm crawling. Forms an extensive filament system that extends
CC from sperm villipoda, along the leading edge of the pseudopod (by
CC similarity).
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24499; AAA29146.1; -.
CC HSSP: P27439; 1MSP.
CC InterPro: IPR000535; MSP_domain.
CC Pfam: PF00635; MSP_domain; 1.
CC Cytokeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13858 MW; 0255C31F187549BC CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8
DB 58 NGVLGP 63
|||||

RESULT 33
MSP3_GLO
ID MSP3_GLO STANDARD; PRT; 125 AA.
AC P53023;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major sperm protein 3.
GN MSP-3.
OS Globodera rostochiensis (Golden nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX NCBI_TaxID=31243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ro1 / Mierenbos;
RA Novitski C.E., Brown S., Chen R., Corner A.S., Atkinson H.J.,
RA McPherson M.J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC Sperm crawling. Forms an extensive filament system that extends
CC from sperm villipoda, along the leading edge of the pseudopod (by
CC similarity).
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24499; AAA29146.1; -.
CC HSSP: P27439; 1MSP.
CC InterPro: IPR000535; MSP_domain.
CC Pfam: PF00635; MSP_domain; 1.
CC Cytokeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13858 MW; 0255C31F187549BC CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8
DB 58 NGVLGP 63
|||||

RESULT 32

```

```

MSP2_GLO
ID MSP2_GLO STANDARD; PRT; 125 AA.
AC P53022;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major sperm protein 2.
GN MSP-2.
OS Globodera rostochiensis (Golden nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX NCBI_TaxID=31243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ro1 / Mierenbos;
RA Novitski C.E., Brown S., Chen R., Corner A.S., Atkinson H.J.,
RA McPherson M.J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC Sperm crawling. Forms an extensive filament system that extends
CC from sperm villipoda, along the leading edge of the pseudopod (by
CC similarity).
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24500; AAA29147.1; -.
CC HSSP: P27439; 1MSP.
CC InterPro: IPR000535; MSP_domain.
CC Pfam: PF00635; MSP_domain; 1.
CC Cytokeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13801 MW; 1B04AD5756511B13 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8
DB 58 NGVLGP 63
|||||

RESULT 33
MSP3_GLO
ID MSP3_GLO STANDARD; PRT; 125 AA.
AC P53023;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major sperm protein 3.
GN MSP-3.
OS Globodera rostochiensis (Golden nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX NCBI_TaxID=31243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ro1 / Mierenbos;
RA Novitski C.E., Brown S., Chen R., Corner A.S., Atkinson H.J.,
RA McPherson M.J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC Sperm crawling. Forms an extensive filament system that extends
CC from sperm villipoda, along the leading edge of the pseudopod (by
CC similarity).
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24500; AAA29147.1; -.
CC HSSP: P27439; 1MSP.
CC InterPro: IPR000535; MSP_domain.
CC Pfam: PF00635; MSP_domain; 1.
CC Cytokeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13801 MW; 1B04AD5756511B13 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8
DB 58 NGVLGP 63
|||||

RESULT 33

```

CC SIMILARITY).

CC -1- TISSUE SPECIFICITY: SPERM.

CC -1- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L24501; AAA29148.1; -

DR HSP; P27439; IMSP.

DR InterPro; IPR000535; MSP_domain.

DR Pfam; PF00635; MSP_domain; 1.

KW Cytokeleton; Acetylation; Sperm; Multigene family.

FT INIT_MET 0 0 BY SIMILARITY.

FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).

SQ SEQUENCE 125 AA; 13844 MW; 0255C5A9C31E49BC CRC64;

Query Match 3.4%; Score 6; DB 1; Length 125;

Best Local Similarity 100.0%; Pred. No. 74;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NGVLGP 8

DB 58 NGVLGP 63

|||||

RESULT 34

FXSA_SERMA STANDARD; PRT; 139 AA.

AC P37148;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fxsa protein (suppressor of F exclusion of phase T7) (Fragment).

GN FXSA.

OS Serratia marcescens.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Serratia.

OX NCBI_TaxID=615;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sr41;

RA MEDLINE=95148727; PubMed=7846149;

RA Omori K., Akatsuka H., Komatsubara S.;

RT "Construction of a versatile promoter analysis vector and its use for

RL analysis of the Serratia marcescens aspartase promoter region.";

RL Plasmid 32:233-237(1994).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Potential).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D13252; BAA02517.1; -

DR PIR; PS0396; PS0396.

KW Transmembrane; Inner membrane.

FT TRANSMEM 17 37 POTENTIAL.

FT TRANSMEM 78 98 POTENTIAL.

FT NON_TER 139 139

SQ SEQUENCE 139 AA; 14996 MW; A897780614A5A09D CRC64;

Query Match 3.4%; Score 6; DB 1; Length 139;

Best Local Similarity 100.0%; Pred. No. 82;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SUTLKL 19

DB 107 SUTLKL 112

|||||

RESULT 35

MYG_PSEGE STANDARD; PRT; 146 AA.

AC Q9DEP1;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Myoglobin.

GN MB.

OS Pseudochaeenichthys georgianus (South Georgia icefish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;

OC Notothenioidae; Channichthyidae; Pseudochaeenichthys.

OX NCBI_TaxID=52239;

RN [1]

RP SEQUENCE FROM N.A.

TISSUE=Heart ventricle;

RX MEDLINE=97430285; PubMed=9284559;

RA Vayda M.E., Small D.J., Yuan M.-L., Costello L., Sidell B.D.;

RT "Conservation of the myoglobin gene among Antarctic notothenioid

RT fishes.";

RL Mol. Mar. Biol. Biotechnol. 6:207-216(1997).

CC -1- FUNCTION: SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES

CC THE MOVEMENT OF OXYGEN WITHIN MUSCLES.

CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U71055; AAG16643.1; -

DR HSP; P02205; IMYT.

DR InterPro; IPR000971; Globin.

DR InterPro; IPR002335; Myoglobin.

DR Pfam; PF00042; globin; 2

DR PRINTS; PR00613; MYOGLOBIN.

DR PROSITE; PS01033; GLOBIN; 1.

KW Heme; Oxygen transport; Transport; Muscle.

FT INIT_MET 0 0 BY SIMILARITY.

FT METAL 59 59 IRON (HEME DISTAL LIGAND).

FT METAL 88 88 IRON (HEME PROXIMAL LIGAND).

SQ SEQUENCE 146 AA; 15574 MW; 8E006AD3EAE0F840 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 146;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVEADY 162

DB 12 LVEADY 17

|||||

RESULT 36

YXIK_BACSU STANDARD; PRT; 153 AA.

ID YXIK_BACSU

AC P42302;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein yxik.

GN YXIK OR SS8B.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 RN NCBI_TaxID=1423;
 [1]
 RC STRAIN=168 / BGSC1A1;
 RX MEDLINE=97124196; PubMed=8969509;
 RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
 RA Miwa Y., Fujita Y.;
 RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
 RT containing the lic and cel loci, and creation of a 177 kb contig
 RT covering the gnt-sacXY region.";
 RL Microbiology 142:3113-3123(1996).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=168;
 RC MEDLINE=98044033; PubMed=9384377;
 RX Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klauer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Oglwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarottl A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D83026; BAAL1690.1; -;
 DR EMBL; 299124; CAB15950.1; -;
 DR Subtilist; BG11140; yxik.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 153 AA; 17649 MW; CEC1799C0BF2A4DC CRC64;
 Query Match 3.4%; Score 6; DB 1; Length 153;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 AGLAYD 98
 Db 39 AGLAYD 44
 I I I I I I I
 RESULT 37
 GRPE_STRMU
 ID ITC_ALOMA STANDARD; PRT; 184 AA.
 AC P35812; 1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin/chymotrypsin inhibitor.
 OS Alocasia macrorrhiza (Giant taro).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Alocasia.
 OX NCBI_TaxID=4456;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. FUI-1;
 RX MEDLINE=94191001; PubMed=8142459;
 RA Argall M.E., Brandbury H.J., Shaw D.C.;
 RT "Amino-acid sequence of a trypsin/chymotrypsin inhibitor from giant
 RT taro (Alocasia macrorrhiza).";
 RL Biochim. Biophys. Acta 1204:189-194(1994).
 CC -!- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR

ID GRPE_STRMU STANDARD; PRT; 180 AA.
 AC O06941;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GRPE protein (HSP-70 cofactor).
 GN GRPE.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=97426041; PubMed=9282745;
 RA Jayaraman G.C., Penders J.E., Burne R.A.;
 RT "Transcriptional analysis of the Streptococcus mutans hrca, grpe and
 RT dna genes and regulation of expression in response to heat shock and
 RT environmental acidification.";
 RL Environmental Microbiol. 25:329-341(1997).
 CC -!- FUNCTION: STIMULATES, JOINTLY WITH DNAJ, THE ATPASE ACTIVITY OF
 CC DNAK. HELPS TO RELEASE ADP FROM DNAK THUS ALLOWING DNAK TO RECYCLE
 CC MORE EFFICIENTLY (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GRPE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U78296; AAC45611.1; -;
 DR HSSP: P09372; 1DKG
 DR InterPro: IPR000740; Grpe.
 DR Pfam: PF01025; Grpe; 1.
 DR PROSITE: PS01071; Grpe; 1.
 KW Chaperone; Heat shock.
 SQ SEQUENCE 180 AA; 20621 MW; 5DD972412CDC5931 CRC64;
 Query Match 3.4%; Score 6; DB 1; Length 180;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 67 KKEEYK 72
 Db 6 KKEEYK 11
 I I I I I I I
 RESULT 38
 ITC_ALOMA STANDARD; PRT; 184 AA.
 AC P35812; 1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin/chymotrypsin inhibitor.
 OS Alocasia macrorrhiza (Giant taro).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Alocasia.
 OX NCBI_TaxID=4456;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. FUI-1;
 RX MEDLINE=94191001; PubMed=8142459;
 RA Argall M.E., Brandbury H.J., Shaw D.C.;
 RT "Amino-acid sequence of a trypsin/chymotrypsin inhibitor from giant
 RT taro (Alocasia macrorrhiza).";
 RL Biochim. Biophys. Acta 1204:189-194(1994).
 CC -!- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR

```
CC FAMILY.
DR PIR: S42753; S42753.
DR HSSP; P07596; IAVA.
DR InterPro: IPR002160; Kunitz_legume.
DR Pfam: PF00157; Kunitz_legume; 1.
DR ProDom: PD000891; Kunitz_legume; 1.
DR SMART; SM00452; STI; 1.
DR PROSITE; PS00283; SOYBEAN_KUNITZ; FALSE_NEG.
KW Serine protease inhibitor.
FT DISULFID 39 84 BY SIMILARITY.
FT DISULFID 136 147 BY SIMILARITY.
FT ACT_SITE 56 57 REACTIVE BOND (POTENTIAL).
FT VARIANT 24 24 M -> A (IN 50% OF THE CHAINS).
FT VARIANT 50 50 E -> K (IN 25% OF THE CHAINS).
SQ SEQUENCE 184 AA; 19774 MW; 60536E9C15F472ED CRC64;

Query Match 3.4%; Score 6; DB 1; Length 184;
Best Local Similarity 100.0%; Pred. No. le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 74 NSRFAT 79
DB 113 NSRFAT 118

RESULT 39
PTH_BACHD STANDARD; PRT; 185 AA.
AC Q9KGJ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peptidyl-tRNA hydrolase (EC 3.1.1.29) (PTH).
GN PTH OR SPOVC OR BH0068.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: THE NATURAL SUBSTRATE FOR THIS ENZYME MAY BE PEPTIDYL-
CC TRNAS WHICH DROP OFF THE RIBOSOME DURING PROTEIN SYNTHESIS (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N-substituted aminoacyl-tRNA + H(2)O -> N-
CC substituted amino acid + tRNA.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PTH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001507; BAB03787.1;
CC HSSP; P23932; 2PTH.
CC InterPro: IPR001328; Pept_trna_hydro.
CC Pfam; PF01195; Pept_trna_hydro; 1.
CC ProDom; PD005324; Pept_trna_hydro; 1.
CC TIGRFAMs; TIGR00447; pth; 1.
CC PROSITE; PS01195; PEPT_TRNA_HYDROL_1; 1.
CC PROSITE; PS01196; PEPT_TRNA_HYDROL_2; 1.
CC Hydrolase; Complete proteome.
```

```
SO SEQUENCE 185 AA; 20547 MW; AAC87F8B5354269D CRC64;

Query Match 3.4%; Score 6; DB 1; Length 185;
Best Local Similarity 100.0%; Pred. No. le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 63 EAFTKK 68
DB 170 EAFTKK 175

RESULT 40
DYL_HSVS7 STANDARD; PRT; 186 AA.
AC P27421;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Dihydrofolate reductase (EC 1.5.1.3) (DHFR).
GN 2 OR DHFR.
OS Herpesvirus saimiri (strain 484-77).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94269832; PubMed=8209420;
RA Geck P., Whitaker S.A., Medveczky M.M., Last T.J., Medveczky P.G.;
RT "Small RNA expression from the oncogenic region of a highly oncogenic
RT strain of herpesvirus saimiri."
RL Virus Genes 8:25-34(1994).
CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) -> 7,8-
CC dihydrofolate + NADPH.
CC -1- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS.
CC DNA PRECURSOR SYNTHESIS AND FOR THE CONVERSION OF DUMP TO DTMP.
CC -1- SIMILARITY: BELONGS TO THE DIHYDROFOLATE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58774; CAA41575.1;
CC PIR; S14840; S14840.
CC HSSP; P00374; IDHF.
CC InterPro: IPR001796; DHFR.
CC Pfam; PF00186; Dihfolate_red; 1.
CC PRINTS; PR00070; DHFR.
CC PROSITE; PS00075; DHFR; 1.
CC Oxidoreductase; NADP: One-carbon metabolism.
CC SEQUENCE 186 AA; 21674 MW; 8249CD3C7ABE1BA2 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 84 SLNDAL 89
DB 92 SLNDAL 97
```

Search completed: May 12, 2003, 09:47:58
Job time : 38 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:44:40 ; Search time 31 Seconds
(without alignments)
1189.756 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 179

Sequence: 1 QHNGVLGPYIGKSLTKLP.....ADYTTKATANLYGLNLNRYR 179

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	7.3	427	2	Q48051 haemophilus
2	13	7.3	449	2	Q9k2q4 haemophilus
3	13	7.3	449	2	Q9khf9 haemophilus
4	13	7.3	449	2	Q00754 haemophilus
5	13	7.3	450	2	Q9khf4 haemophilus
6	13	7.3	450	2	Q9khf3 haemophilus
7	13	7.3	451	2	Q9khf2 haemophilus
8	13	7.3	451	2	Q9khg0 haemophilus
9	13	7.3	451	2	Q9khf8 haemophilus
10	13	7.3	451	2	Q9khf7 haemophilus
11	13	7.3	451	2	Q9khf6 haemophilus
12	13	7.3	454	2	Q9khg1 haemophilus
13	13	7.3	455	2	Q9khH5 haemophilus
14	13	7.3	455	2	Q9khH4 haemophilus
15	13	7.3	455	2	Q9khH3 haemophilus
16	13	7.3	455	2	Q9khH2 haemophilus

Q9khh1	haemophilus	17	13	7.3	455	2	Q9KHH1
Q00450	haemophilus	18	13	7.3	455	2	Q00450
Q9khi4	haemophilus	19	13	7.3	456	2	Q9KH14
Q9khi3	haemophilus	20	13	7.3	456	2	Q9KH13
Q9khi2	haemophilus	21	13	7.3	456	2	Q9KH12
Q9khi1	haemophilus	22	13	7.3	456	2	Q9KH11
Q9khi0	haemophilus	23	13	7.3	456	2	Q9KH10
Q9khh9	haemophilus	24	13	7.3	456	2	Q9KHH9
Q9khh6	haemophilus	25	13	7.3	456	2	Q9KHH6
Q9khh9	haemophilus	26	13	7.3	456	2	Q9KHG9
Q9khh8	haemophilus	27	13	7.3	456	2	Q9KHG8
Q9khh7	haemophilus	28	13	7.3	456	2	Q9KHG7
Q9khh5	haemophilus	29	13	7.3	456	2	Q9KHG5
Q9khh4	haemophilus	30	13	7.3	456	2	Q9KHG4
Q9khh3	haemophilus	31	13	7.3	456	2	Q9KHG3
Q9khh6	haemophilus	32	13	7.3	457	2	Q9KHG6
Q9k397	haemophilus	33	13	7.3	458	2	Q9K397
Q9k324	haemophilus	34	13	7.3	459	2	Q9K324
Q9khh8	haemophilus	35	13	7.3	459	2	Q9KHH8
Q9khh7	haemophilus	36	13	7.3	459	2	Q9KHH7
Q9khh0	haemophilus	37	13	7.3	460	2	Q9KHH0
Q9khf5	haemophilus	38	10	5.6	449	2	Q9KHf5
Q9k184	vibrlo chol	39	9	5.0	412	16	Q9KL84
O28679	archaeoglob	40	8	4.5	72	17	O28679
Q8xdd4	escherichia	41	8	4.5	417	16	Q8XDD4
Q8v972	lily mottle	42	8	4.5	622	12	Q8V972
Q8ulh0	pyrococcus	43	8	4.5	632	17	Q8ULH0
Q9v0x7	pyrococcus	44	8	4.5	633	17	Q9V0X7
Q9n4k4	caenorhabdl	45	8	4.5	649	5	Q9N4K4
O8uyt2	lily mottle	46	8	4.5	759	12	O8UYT2
Q96xb5	sulfolobus	47	7	3.9	59	17	Q96XB5
Q9hw3	schizosacch	48	7	3.9	68	3	Q9HWD3
Q9d4s9	mus musculus	49	7	3.9	127	11	Q9D4S9
Q9fkf5	arabidopsis	50	7	3.9	134	10	Q9FKF5
Q9l0g5	streptomyce	51	7	3.9	144	16	Q9L0G5
Q9g12	thunnus ala	52	7	3.9	147	13	Q9DGJ2
Q9g17	thunnus alb	53	7	3.9	147	13	Q9DGJ7
Q8uda8	agrobacteri	54	7	3.9	180	16	Q8UDA8
Q937m6	photorhabdu	55	7	3.9	184	2	Q937M6
Q9c9s5	arabidopsis	56	7	3.9	185	10	Q9C9S5
Q949x6	arabidopsis	57	7	3.9	193	10	Q949X6
Q8vwal	streptomyce	58	7	3.9	213	2	Q8VWA1
P93007	arabidopsis	59	7	3.9	218	10	P93007
Q9eww7	streptomyce	60	7	3.9	220	16	Q9EWW7
O21l06	arabidopsis	61	7	3.9	229	10	O21L06
Q40479	nicotiana t	62	7	3.9	233	10	Q40479
Q9lw50	nicotiana s	63	7	3.9	237	10	Q9LW50
Q9lkk0	atriplex ho	64	7	3.9	240	10	Q9LKK0
Q9dcy9	mus musculus	65	7	3.9	260	11	Q9DCY9
Q9lmi5	arabidopsis	66	7	3.9	261	10	Q9LMI5
Q9tc83	centropomus	67	7	3.9	274	8	Q9TC83
Q9zkr3	helicobacte	68	7	3.9	274	16	Q9ZKR3
O8vfg3	mus musculus	69	7	3.9	314	11	O8VFG3
Q9sqz3	arabidopsis	70	7	3.9	330	10	Q9SGZ3
Q9ak95	arabidopsis	71	7	3.9	334	10	Q9AK95
O8tuc2	methanosarc	72	7	3.9	345	17	O8TUC2
Q9gn06	babesia big	73	7	3.9	351	5	Q9GN06
Q9hsk5	halobacteri	74	7	3.9	363	17	Q9HSK5
O9l102	streptomyce	75	7	3.9	413	16	O9L102
O69679	mycobacteri	76	7	3.9	413	16	O69679
O8zd48	yersinia pe	77	7	3.9	423	16	O8ZD48
Q8zna5	salmonella	78	7	3.9	437	16	Q8ZNA5
O8z4y8	salmonella	79	7	3.9	437	16	O8Z4Y8
O8xcn6	escherichia	80	7	3.9	448	16	O8XCN6
O9l9m4	oryza sativ	81	7	3.9	488	10	O9L9M4
O8zx57	pyrobaculum	82	7	3.9	499	17	O8ZX57
O25l25	helicobacte	83	7	3.9	511	16	O25L25
Q9zkb7	helicobacte	84	7	3.9	511	16	Q9ZKB7
Q44372	agrobacteri	85	7	3.9	526	10	Q44372
Q9szq3	arabidopsis	86	7	3.9	526	2	Q9SZQ3
O8cm3	yersinia pe	87	7	3.9	528	16	O8CM3
Q9las0	desulfotoba	88	7	3.9	536	2	Q9LAS0
Q8ulp0	pyrococcus	89	7	3.9	551	17	Q8ULP0

236	6	3.4	200	17	Q97B00	Q97b00 thermoplasm	309	6	3.4	247	2	Q00442	Q00442 saccharopol
237	6	3.4	201	2	P96514	P96514 borrelia ja	310	6	3.4	247	4	Q9H3H0	Q9H3H0 homo sapien
238	6	3.4	201	2	O08234	O08234 borrelia ta	311	6	3.4	247	17	Q9HQA4	Q9HQA4 halobacteri
239	6	3.4	201	2	P96516	P96516 borrelia ja	312	6	3.4	250	16	Q25124	Q25124 helicobacte
240	6	3.4	201	2	P96572	P96572 borrelia ja	313	6	3.4	250	16	Q9ZKB6	Q9ZKB6 helicobacte
241	6	3.4	201	2	P96573	P96573 borrelia ja	314	6	3.4	251	17	Q8TJY4	Q8TJY4 methanosarc
242	6	3.4	202	2	P96508	P96508 borrelia af	315	6	3.4	252	5	Q8T227	Q8T227 dictyostell
243	6	3.4	202	2	O50626	O50626 borrelia af	316	6	3.4	253	3	Q8WZQ8	Q8WZQ8 neurospora
244	6	3.4	202	2	Q93Q97	Q93Q97 borrelia bu	317	6	3.4	253	5	Q9V5T6	Q9V5T6 drosophila
245	6	3.4	203	2	O50619	O50619 borrelia bu	318	6	3.4	253	16	O50988	O50988 borrelia bu
246	6	3.4	203	2	O50624	O50624 borrelia af	319	6	3.4	253	16	Q9HQA4	Q9HQA4 pseudomonas
247	6	3.4	203	2	O44985	O44985 borrelia ga	320	6	3.4	254	16	Q92613	Q92613 listeria in
248	6	3.4	203	10	O82447	O82447 nicotiana t	321	6	3.4	254	16	Q8YI08	Q8YI08 brucella me
249	6	3.4	204	2	P96509	P96509 borrelia af	322	6	3.4	256	17	O8TYA8	O8TYA8 methanopyru
250	6	3.4	204	2	O44989	O44989 borrelia af	323	6	3.4	257	5	O01336	O01336 caenorhabdi
251	6	3.4	204	16	Q9WYW1	Q9WYW1 thermotoga	324	6	3.4	257	10	Q940K8	Q940K8 arabidopsals
252	6	3.4	204	16	Q9A298	Q9A298 caulobacter	325	6	3.4	258	16	Q54544	Q54544 synecocyst
253	6	3.4	205	17	Q8TVN0	Q8TVN0 methanopyru	326	6	3.4	259	10	O65544	O65544 arabidopsals
254	6	3.4	206	16	Q9JQN8	Q9JQN8 neisseria m	327	6	3.4	260	16	Q92VL9	Q92VL9 rhizobium m
255	6	3.4	207	2	Q45175	Q45175 borrelia ga	328	6	3.4	261	2	O8RP05	O8RP05 azospirillum
256	6	3.4	207	10	Q96230	Q96230 allocasia ma	329	6	3.4	261	5	Q9V3C2	Q9V3C2 drosophila
257	6	3.4	208	16	O8XCL3	O8XCL3 escherichia	330	6	3.4	261	16	Q986Y9	Q986Y9 rhizobium l
258	6	3.4	208	16	Q980P6	Q980P6 mycoplasma	331	6	3.4	261	16	O54172	O54172 streptomyce
259	6	3.4	209	2	Q44671	Q44671 borrelia af	332	6	3.4	261	16	Q9WXQ6	Q9WXQ6 thermotoga
260	6	3.4	209	2	Q9579	Q9579 borrelia ga	333	6	3.4	263	2	Q9RB25	Q9RB25 pectobacter
261	6	3.4	210	2	Q9EVT8	Q9EVT8 thermus the	334	6	3.4	263	2	Q9RB24	Q9RB24 pectobacter
262	6	3.4	210	2	Q92679	Q92679 borrelia sp	335	6	3.4	263	2	Q9RB21	Q9RB21 pectobacter
263	6	3.4	210	2	Q9F435	Q9F435 klebsiella	336	6	3.4	263	2	Q9RB23	Q9RB23 pectobacter
264	6	3.4	210	5	O9N6L5	O9N6L5 leishmania	337	6	3.4	263	2	Q9X4A8	Q9X4A8 pectobacter
265	6	3.4	211	2	Q9576	Q9576 borrelia af	338	6	3.4	263	5	O16722	O16722 caenorhabdi
266	6	3.4	212	2	Q9KIM5	Q9KIM5 borrelia af	339	6	3.4	263	16	Q8ZNF7	Q8ZNF7 salmonella
267	6	3.4	212	10	Q9STQ4	Q9STQ4 arabidopsals	340	6	3.4	263	16	Q8ZFI9	Q8ZFI9 versinia pe
268	6	3.4	213	17	O8U2V0	O8U2V0 pyrococcus	341	6	3.4	263	16	O82667	O82667 salmonella
269	6	3.4	215	10	Q9F212	Q9F212 solanum tub	342	6	3.4	263	16	O8XCN4	O8XCN4 escherichia
270	6	3.4	216	2	Q9RN47	Q9RN47 salmonella	343	6	3.4	264	16	Q8YFI0	Q8YFI0 brucella me
271	6	3.4	217	3	Q20335	Q20335 saccharomyc	344	6	3.4	265	16	Q98BU4	Q98BU4 rhizobium l
272	6	3.4	217	4	Q96BD2	Q96BD2 homo sapien	345	6	3.4	266	12	O72160	O72160 southern be
273	6	3.4	219	4	Q96PJ4	Q96PJ4 homo sapien	346	6	3.4	266	2	O51673	O51673 mycobacterl
274	6	3.4	219	4	Q9NTL0	Q9NTL0 homo sapien	347	6	3.4	267	2	Q51673	Q51673 pyrococcus
275	6	3.4	221	10	Q9R1L2	Q9R1L2 hemerocalli	348	6	3.4	267	16	Q97RN2	Q97RN2 streptococc
276	6	3.4	224	11	O8RIW1	O8RIW1 mus musculu	349	6	3.4	267	16	Q8YI89	Q8YI89 ralistonia s
277	6	3.4	225	4	Q96PJ1	Q96PJ1 homo sapien	350	6	3.4	268	10	O90277	O90277 lycopersico
278	6	3.4	226	4	Q96A03	Q96A03 homo sapien	351	6	3.4	268	16	Q9A034	Q9A034 streptococc
279	6	3.4	226	11	Q9CQC3	Q9CQC3 mus musculu	352	6	3.4	269	16	Q97RA2	Q97RA2 streptococc
280	6	3.4	228	16	O25181	O25181 helicobacte	353	6	3.4	269	17	O970N6	O970N6 sulfolobus
281	6	3.4	228	16	O25819	O25819 helicobacte	354	6	3.4	270	3	O13758	O13758 schizosacch
282	6	3.4	228	16	Q9ZK06	Q9ZK06 helicobacte	355	6	3.4	270	4	Q96PJ2	Q96PJ2 homo sapien
283	6	3.4	229	16	Q9HT77	Q9HT77 pseudomonas	356	6	3.4	270	5	Q9VKA9	Q9VKA9 drosophila
284	6	3.4	230	5	O17452	O17452 drosophila	357	6	3.4	272	10	O06765	O06765 glycine max
285	6	3.4	230	17	O8THH4	O8THH4 methanosarc	358	6	3.4	272	16	Q989S2	Q989S2 rhizobium l
286	6	3.4	231	4	O8WUI9	O8WUI9 homo sapien	359	6	3.4	273	4	Q92951	Q92951 homo sapien
287	6	3.4	231	16	Q9RJB6	Q9RJB6 streptomyce	360	6	3.4	273	10	Q9SS72	Q9SS72 arabidopsals
288	6	3.4	233	5	Q95YJ3	Q95YJ3 samia cynth	361	6	3.4	273	10	Q8SB81	Q8SB81 oryza sativ
289	6	3.4	233	5	Q95YJ2	Q95YJ2 samia cynth	362	6	3.4	273	13	O57659	O57659 gallus gall
290	6	3.4	233	16	Q985R3	Q985R3 rhizobium l	363	6	3.4	275	5	Q8SZ85	Q8SZ85 drosophila
291	6	3.4	233	17	Q975G6	Q975G6 sulfolobus	364	6	3.4	277	17	Q9HQC4	Q9HQC4 halobacteri
292	6	3.4	233	17	Q96Y77	Q96Y77 sulfolobus	365	6	3.4	278	10	Q948J9	Q948J9 arabidopsals
293	6	3.4	234	10	O04680	O04680 lycopersico	366	6	3.4	278	16	Q8Cmv0	Q8Cmv0 methanosarc
294	6	3.4	234	16	Q9RYI8	Q9RYI8 deinococcus	367	6	3.4	279	16	Q98R68	Q98R68 mycoplasma
295	6	3.4	236	5	Q9VI57	Q9VI57 drosophila	368	6	3.4	279	16	Q97KB0	Q97KB0 clostridium
296	6	3.4	236	10	O40476	O40476 nicotiana t	369	6	3.4	280	5	Q9VKB8	Q9VKB8 ralistonia s
297	6	3.4	236	16	O8YQH0	O8YQH0 anabaena sp	370	6	3.4	280	17	Q8ZSM1	Q8ZSM1 pyrobaculum
298	6	3.4	236	16	O8XLC1	O8XLC1 clostridium	371	6	3.4	281	2	Q9EY14	Q9EY14 streptomyce
299	6	3.4	239	16	Q9A125	Q9A125 streptococc	372	6	3.4	281	4	O8TBP2	O8TBP2 homo sapien
300	6	3.4	239	16	Q9S233	Q9S233 streptomyce	373	6	3.4	281	10	Q9SN47	Q9SN47 arabidopsals
301	6	3.4	240	2	Q9ADY8	Q9ADY8 agrobacteri	374	6	3.4	281	10	O8S015	O8S015 oryza sativ
302	6	3.4	241	5	Q9N4P9	Q9N4P9 caenorhabdi	375	6	3.4	281	16	Q92B68	Q92B68 streptomyce
303	6	3.4	241	17	Q8TWK2	Q8TWK2 methanopyru	376	6	3.4	283	10	Q9FJ53	Q9FJ53 arabidopsals
304	6	3.4	243	4	Q9H6F1	Q9H6F1 homo sapien	377	6	3.4	283	16	Q8X512	Q8X512 escherichia
305	6	3.4	243	16	Q93J75	Q93J75 streptomyce	378	6	3.4	283	16	Q8U8Y4	Q8U8Y4 agrobacteri
306	6	3.4	244	16	O8YDQ9	O8YDQ9 brucella me	379	6	3.4				
307	6	3.4	245	16	Q9RV09	Q9RV09 deinococcus	380	6	3.4				
308	6	3.4	247	2	Q9JN61	Q9JN61 saccharopol	381	6	3.4				

528	6	3.4	360	13	Q9YGC9	O9ygc9 brachydanio	601	403	10	O23109	O23109 arabldopsls
529	6	3.4	362	4	Q96S93	Q96s93 homo sapien	602	404	16	O85338	O85338 escherichia
530	6	3.4	362	5	Q22897	Q22897 caenorhabdi	603	405	4	O75295	O75295 homo sapien
531	6	3.4	362	16	O8XVM7	O8xvm7 ralstonia s	604	405	11	Q9CZS7	Q9czs7 mus musculus
532	6	3.4	363	2	O07652	O07652 cellvibrio	605	406	4	Q9NMU3	Q9nmw3 homo sapien
533	6	3.4	363	2	O52284	O52284 agrobacteri	606	406	6	O13010	O13010 sus scrofa
534	6	3.4	363	16	O84198	O84198 chlamydia t	607	406	11	Q9R0I8	Q9rdj2 rattus norv
535	6	3.4	365	16	O9KFJ4	O9kfj4 bacillus ha	608	407	16	Q9ZDJ2	Q9zjd2 rickettsia
536	6	3.4	365	16	O98P16	O98p16 rhizobium l	609	407	16	O84558	O84558 chlamydia t
537	6	3.4	366	10	O9LYE7	O9lye7 arabldopsls	610	411	16	O84558	O84558 chlamydia t
538	6	3.4	367	10	O9XHC9	O9xhc9 hyacinthus	611	412	5	Q965R2	Q965r2 caenorhabdi
539	6	3.4	367	11	O09179	O09179 rattus norv	612	412	11	Q9C083	Q9cu83 mus musculus
540	6	3.4	368	2	Q93QC6	Q93qc6 corynebacte	613	413	16	O8UF99	O8uf99 agrobacteri
541	6	3.4	368	11	O8R1X0	O8r1x0 mus musculus	614	414	10	Q9C8X8	Q9c8x8 arabldopsls
542	6	3.4	369	16	O9R2U0	O9r2u0 deinococcus	615	415	10	Q9SY27	Q9sy27 arabldopsls
543	6	3.4	370	16	O9KVS3	O9kvs3 vibrio chol	616	416	4	P78356	P78356 homo sapien
544	6	3.4	371	16	P95291	P95291 mycobacteri	617	416	11	O88377	O88377 rattus norv
545	6	3.4	374	5	O76950	O76950 helobdella	618	417	2	Q9RPA9	Q9rpa9 streptomyce
546	6	3.4	374	6	Q9TU81	Q9tu81 ovis aries	619	418	10	O9LKW7	O9lkw7 zea mays (m
547	6	3.4	374	10	O8RUR1	O8rur1 cryptomeria	620	418	10	O944N8	O944n8 arabldopsls
548	6	3.4	376	4	O8WXH3	O8wxh3 homo sapien	621	418	16	Q9HWP4	Q9hwp4 pseudomonas
549	6	3.4	377	10	O82563	O82563 anemia phyl	622	419	13	O8QG81	O8qg81 xenopus lae
550	6	3.4	377	10	Q9M1P0	Q9m1p0 arabldopsls	623	420	11	O88370	O88370 rattus norv
551	6	3.4	378	2	Q9EUS5	Q9eus5 streptomyce	624	420	11	O81635	O81635 mus musculus
552	6	3.4	379	10	Q9C9Q8	Q9c9q8 arabldopsls	625	420	11	Q99M59	Q99m59 mus musculus
553	6	3.4	379	16	O9AK94	O9ak94 streptomyce	626	420	16	Q8XMD3	Q8xmd3 clostridium
554	6	3.4	381	16	O8Y316	O8y316 ralstonia s	627	421	4	O8TDX8	O8tdx8 homo sapien
555	6	3.4	382	16	O99YV5	O99yv5 streptococc	628	421	11	O91XU3	O91xu3 mus musculus
556	6	3.4	383	2	Q9F5B6	Q9f5b6 agrobacteri	629	422	16	O8YK21	O8yk21 anabaena sp
557	6	3.4	383	10	O9LEL8	O9lel8 citrus sine	630	422	8	Q988Y7	Q988y7 fasciola he
558	6	3.4	383	16	Q9CG58	Q9cg58 lactococcus	631	424	2	O50990	O50990 nelisseria 9
559	6	3.4	383	16	Q9K2X4	Q9k2x4 streptomyce	632	424	4	Q9UF27	Q9uf27 homo sapien
560	6	3.4	385	10	O9FNA7	O9fna7 arabldopsls	633	424	16	Q9K0C1	Q9k0c1 nelisseria m
561	6	3.4	385	10	O93X83	O93x83 vitis labru	634	424	16	O9JVC6	O9jvc6 nelisseria m
562	6	3.4	386	16	O97K26	O97k26 clostridium	635	426	17	O972L9	O972l9 sulfobus
563	6	3.4	386	16	O8ZHT6	O8zht6 yersinia pe	636	427	10	O9CAX5	O9cax5 arabldopsls
564	6	3.4	386	16	O8YBH9	O8ybh9 brucella me	637	428	10	Q93WA3	Q93wa3 oryza sativ
565	6	3.4	388	16	O9HW46	O9hw46 pseudomonas	638	428	16	O9A548	O9a548 caulobacter
566	6	3.4	388	16	Q9RY22	Q9ry22 deinococcus	639	429	16	O887R8	O887r8 rhizobium l
567	6	3.4	388	16	Q92KC0	Q92kc0 rhizobium m	640	429	16	O8UEN8	O8uen8 agrobacteri
568	6	3.4	388	16	O8XV11	O8xv11 ralstonia s	641	429	16	O92R78	O92r78 rhizobium m
569	6	3.4	388	17	O8U056	O8u056 pyrococcus	642	430	16	Q9XAQ1	Q9xaq1 streptomyce
570	6	3.4	389	16	O98118	O98118 rhizobium l	643	430	16	Q21891	Q21891 caenorhabdi
571	6	3.4	389	16	O8RHE9	O8rhe9 fusobacteri	644	431	5	Q95KG7	Q95kg7 macaca fasc
572	6	3.4	390	2	Q92ID8	Q92id8 actinobacil	645	431	16	Q9PFB0	Q9pfb0 xyliella fas
573	6	3.4	390	13	P79753	P79753 fugu rubrip	646	432	2	Q9X592	Q9x592 azospiriliu
574	6	3.4	391	12	O80871	O80871 non-a non-b	647	432	16	Q9KT63	Q9kt63 vibrio chol
575	6	3.4	391	16	P72824	P72824 synecocyst	648	432	16	Q9A6J9	Q9a6j9 caulobacter
576	6	3.4	391	16	Q92721	Q92721 listeria ln	649	432	16	Q9PPV2	Q9ppv2 ureaplasma
577	6	3.4	391	16	O8ZRM5	O8zrm5 salmonella	650	434	12	Q9YVR1	Q9yvr1 melanoplus
578	6	3.4	391	16	O8Z986	O8z986 salmonella	651	434	16	O913J2	O913j2 pseudomonas
579	6	3.4	391	17	Q97ZAT	Q97za7 sulfobus	652	434	16	O8RFR6	O8rf6 fusobacteri
580	6	3.4	393	4	Q9BY52	Q9by52 homo sapien	653	435	10	Q94FR3	Q94fr3 triticum ae
581	6	3.4	393	4	Q96SK0	Q96sk0 homo sapien	654	437	5	Q9BHB6	Q9bhb6 plasmodium
582	6	3.4	393	4	O9BWS9	Q9bws9 homo sapien	655	437	5	Q9BHB4	Q9bhb4 plasmodium
583	6	3.4	394	2	O8RNS5	O8rns5 corynebacte	656	437	5	Q9BH76	Q9bh76 plasmodium
584	6	3.4	394	4	O9H836	O9h836 homo sapien	657	437	5	Q9BH63	Q9bh63 plasmodium
585	6	3.4	394	5	O8SVG5	O8svg5 encephalito	658	437	5	Q9BH26	Q9bh26 plasmodium
586	6	3.4	394	10	Q9M0Y4	Q9m0y4 arabldopsls	659	437	5	Q9BHS1	Q9bhs1 plasmodium
587	6	3.4	396	5	Q8SZ14	Q8sz14 drosophila	660	437	5	Q9BHS0	Q9bhs0 plasmodium
588	6	3.4	396	16	O34367	O34367 bacillus su	661	437	5	Q9BHS9	Q9bhs9 plasmodium
589	6	3.4	397	16	O9RYF1	O9ryf1 deinococcus	662	437	5	Q9BHR8	Q9bhr8 plasmodium
590	6	3.4	397	17	O8TLL1	O8tll1 methanosarc	663	437	5	Q9BHR7	Q9bhr7 plasmodium
591	6	3.4	398	11	O9D8A3	O9d8a3 mus musculus	664	437	5	Q9BHR6	Q9bhr6 plasmodium
592	6	3.4	398	11	O8VCX9	O8vcx9 mus musculus	665	437	5	Q9BHR5	Q9bhr5 plasmodium
593	6	3.4	400	16	O8ZPBS	O8zpbs salmonella	666	437	5	Q9BHR4	Q9bhr4 plasmodium
594	6	3.4	400	16	O8Z763	O8z763 salmonella	667	437	5	Q9BHR3	Q9bhr3 plasmodium
595	6	3.4	400	16	O8Y091	O8y091 ralstonia s	668	437	5	Q9BHR2	Q9bhr2 plasmodium
596	6	3.4	401	17	O9HNB5	O9hnb5 halobacteri	669	437	5	Q9BHR1	Q9bhr1 plasmodium
597	6	3.4	401	17	O975F9	O975f9 sulfobus	670	437	5	Q9BHR0	Q9bhr0 plasmodium
598	6	3.4	402	5	Q9BIM8	Q9bim8 plasmodium	671	437	5	Q9BHQ9	Q9bhq9 plasmodium
599	6	3.4	403	2	Q9AE01	Q9ae01 amycolatops	672	437	5	Q9BHQ8	Q9bhq8 plasmodium
600	6	3.4	403	2	Q9XDW7	Q9xdw7 rhodopseudo	673	437	5	Q9BHQ7	Q9bhq7 plasmodium


```
966 6 3.4 622 5 0967J1 Q967J1 plasmodium
967 6 3.4 622 5 0967J0 Q967J0 plasmodium
968 6 3.4 622 5 0967I9 Q967I9 plasmodium
969 6 3.4 622 5 0967I8 Q967I8 plasmodium
970 6 3.4 622 5 0967I7 Q967I7 plasmodium
971 6 3.4 622 5 0967I6 Q967I6 plasmodium
972 6 3.4 622 5 0967I5 Q967I5 plasmodium
973 6 3.4 622 5 0967I4 Q967I4 plasmodium
974 6 3.4 622 5 0967I3 Q967I3 plasmodium
975 6 3.4 622 5 0967I2 Q967I2 plasmodium
976 6 3.4 622 5 095N21 Q95N21 plasmodium
977 6 3.4 622 5 095N04 Q95N04 plasmodium
978 6 3.4 622 5 095N02 Q95N02 plasmodium
979 6 3.4 622 5 094661 Q94661 plasmodium
980 6 3.4 622 5 09H967 Q9H967 homo sapien
981 6 3.4 627 10 Q9SL48 Q9SL48 arabisopsis
982 6 3.4 629 4 Q95671 Q95671 homo sapien
983 6 3.4 629 10 Q94GH1 Q94GH1 oryza sativ
984 6 3.4 631 5 Q45873 Q45873 caenorhabdi
985 6 3.4 633 10 Q95I73 Q95I73 arabisopsis
986 6 3.4 633 10 Q9FV33 Q9FV33 picea abies
987 6 3.4 633 17 Q95I76 Q95I76 pyrococcus
988 6 3.4 633 17 Q9HP10 Q9HP10 halobacteri
989 6 3.4 634 5 Q22585 Q22585 caenorhabdi
990 6 3.4 636 2 Q938A9 Q938A9 pseudomonas
991 6 3.4 642 4 Q9P263 Q9P263 homo sapien
992 6 3.4 645 8 Q63620 Q63620 balanogloss
993 6 3.4 645 16 Q9PNH8 Q9PNH8 campylobact
994 6 3.4 645 16 Q9RDB4 Q9RDB4 streptomyce
995 6 3.4 645 16 Q9PWT3 Q9PWT3 streptomyce
996 6 3.4 650 2 Q8RR67 Q8RR67 actinobacil
997 6 3.4 653 10 Q9SS77 Q9SS77 arabisopsis
998 6 3.4 653 16 Q9I2I2 Q9I2I2 pseudomonas
999 6 3.4 657 10 Q9LFX6 Q9LFX6 arabisopsis
1000 6 3.4 658 4 Q9BQ26 Q9BQ26 homo sapien
```

ALIGNMENTS

```
RESULT 1
Q48051 ID Q48051 PRELIMINARY; PRT; 427 AA.
AC Q48051;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Outer membrane protein P1 (Fragment).
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF3029;
RX MEDLINE=96355877; PubMed=8751915;
RA "RNA sequence analysis and restriction fragment length polymorphisms
RT of the P1 gene of Haemophilus influenzae biogroup aegyptius associated
RT with Brazilian purpuric fever."
RL Infect. Immun. 64:3666-3672(1996).
DR EMBL; U28808; AAC44459.1; -
FT NON_TER 1 1
FT VARIANT 61 61 D -> Y.
FT VARIANT 265 265 D -> N.
FT VARIANT 385 385 Y -> H.
SQ SEQUENCE 427 AA; 46489 MW; 80620B4A6FA0DD27 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
Db 390 GATYKFTPNLSVD 402

RESULT 3
Q9KHF9 ID Q9KHF9 PRELIMINARY; PRT; 449 AA.
AC Q9KHF9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-105;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc O.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260370; AAF97584.1; -
DR EMBL; AF260368; AAF97582.1; -
SQ SEQUENCE 449 AA; 48526 MW; 693B515B2112B536 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
Db 390 GATYKFTPNLSVD 402

RESULT 3
Q9KHF9 ID Q9KHF9 PRELIMINARY; PRT; 449 AA.
AC Q9KHF9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-105;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc O.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260370; AAF97584.1; -
DR EMBL; AF260368; AAF97582.1; -
SQ SEQUENCE 449 AA; 48526 MW; 693B515B2112B536 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
Db 390 GATYKFTPNLSVD 402
```

```

RESULT 4
Q00754
ID Q00754 PRELIMINARY; PRT; 449 AA.
AC Q00754;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Outer membrane protein P1 precursor (OMP P1).
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90035394; PubMed=2572549;
RA Munson R.S. Jr., Grass S., Einhorn M., Bailey C., Newell C.;
RT "Comparative analysis of the structures of the outer membrane protein
RT P1 genes from major clones of Haemophilus influenzae type b.";
RL Infect. Immun. 57:3300-3305(1989).
DR EMBL: W73494; AAA24992.1; ...
KW Outer membrane; Signal; Transmembrane.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 449 OUTER MEMBRANE PROTEIN P1.
SQ SEQUENCE 449 AA; 48315 MW; 29C668F90AECDF9 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 390 GATYKFTPNLSVD 402
|||||
|||||

RESULT 5
Q0KHf4
ID Q0KHf4 PRELIMINARY; PRT; 450 AA.
AC Q0KHf4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=7424;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260375; AAF97589.1; ...
KW Infect. Immun. 68:4505-4517(2000).
SQ SEQUENCE 450 AA; 48593 MW; CF17201C42BDA757 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 392 GATYKFTPNLSVD 404
|||||
|||||

RESULT 6
Q0KHf3
ID Q0KHf3 PRELIMINARY; PRT; 450 AA.
AC Q0KHf3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=4-H-1093;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260376; AAF97590.1; ...
KW Infect. Immun. 68:4505-4517(2000).
SQ SEQUENCE 450 AA; 48741 MW; 3C68E144339F41CD CRC64;

Query Match 7.3%; Score 13; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 392 GATYKFTPNLSVD 404
|||||
|||||

RESULT 7
Q0KHG2
ID Q0KHG2 PRELIMINARY; PRT; 451 AA.
AC Q0KHG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=6181;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260365; AAF97579.1; ...
KW Infect. Immun. 68:4505-4517(2000).
SQ SEQUENCE 451 AA; 48611 MW; B616AAB6FEF0252A CRC64;

Query Match 7.3%; Score 13; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 391 GATYKFTPNLSVD 403
|||||
|||||

RESULT 8
Q0KHG0
ID Q0KHG0 PRELIMINARY; PRT; 451 AA.
AC Q0KHG0;

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
GN Outer membrane protein P1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=4-H-1094;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260367; AAF97581.1; -; 72C6977C6F2A7BA1 CRC64;
SQ SEQUENCE 451 AA; 48640 MW; 72C6977C6F2A7BA1 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 391 GATYKFTPNLSVD 403

RESULT 9
Q9KHF8 PRELIMINARY; PRT; 451 AA.
ID Q9KHF8
AC Q9KHF8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=2-H-1038;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260371; AAF97585.1; -;
SQ SEQUENCE 451 AA; 48631 MW; E9D7DF99390C0BDB CRC64;

Query Match 7.3%; Score 13; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 392 GATYKFTPNLSVD 404

RESULT 10
Q9KHF7 PRELIMINARY; PRT; 451 AA.
ID Q9KHF7
AC Q9KHF7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

```

```

DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=658;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260372; AAF97586.1; -; 3C6631D057EB0FFB CRC64;
SQ SEQUENCE 451 AA; 48796 MW; 3C6631D057EB0FFB CRC64;

Query Match 7.3%; Score 13; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 392 GATYKFTPNLSVD 404

RESULT 11
Q9KHF6 PRELIMINARY; PRT; 451 AA.
ID Q9KHF6
AC Q9KHF6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=13-H-1157;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260373; AAF97587.1; -;
SQ SEQUENCE 451 AA; 48446 MW; F4DD7C9354DFCF13 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 392 GATYKFTPNLSVD 404

RESULT 12
Q9KHG1 PRELIMINARY; PRT; 454 AA.
ID Q9KHG1
AC Q9KHG1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.

```



```

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=22-H-1154;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260366; AAF97380.1; -.
SQ SEQUENCE 454 AA; 48986 MW; 494FBAC2600DA2C8 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|
Db 392 GATYKFTPNLSVD 404

RESULT 13
Q9KHH5 PRELIMINARY; PRT; 455 AA.
AC Q9KHH5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=7004;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260348; AAF97362.1; -.
SQ SEQUENCE 455 AA; 49165 MW; 280E211B16CA18C0 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|
Db 393 GATYKFTPNLSVD 405

RESULT 14
Q9KHH4 PRELIMINARY; PRT; 455 AA.
AC Q9KHH4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=6094;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260349; AAF97563.1; -.
SQ SEQUENCE 455 AA; 49111 MW; 0CB00B162A6A6CBI CRC64;

Query Match 7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|
Db 393 GATYKFTPNLSVD 405

RESULT 15
Q9KHH3 PRELIMINARY; PRT; 455 AA.
AC Q9KHH3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC9795;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL: AF260352; AAF97566.1; -.
SQ SEQUENCE 455 AA; 49169 MW; 77A8BA97BCE2B450 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|
Db 393 GATYKFTPNLSVD 405

RESULT 16
Q9KHH2 PRELIMINARY; PRT; 455 AA.
AC Q9KHH2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=161;

```

```

RX MEDLINE-20359342; PubMed-10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260353; AAF97567.1; -.
SQ SEQUENCE 455 AA; 49139 MW; 9CCB2EE675B421D9 CRC64;

Query Match          7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 393 GATYKFTPNLSVD 405

RESULT 17
ID Q9KHH1 PRELIMINARY; PRT; 455 AA.
AC Q9KHH1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OMPPI.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID-727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OT-9;
RX MEDLINE-20359342; PubMed-10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260354; AAF97568.1; -.
SQ SEQUENCE 455 AA; 49123 MW; FD5882979F10EB24 CRC64;

Query Match          7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 393 GATYKFTPNLSVD 405

RESULT 18
Q00450 PRELIMINARY; PRT; 455 AA.
ID Q00450;
AC Q00450;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Outer membrane protein PI precursor, subtype 3L (OMP PI).
GN OMPPI.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID-727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1613 / SEROTYPE B;
RX MEDLINE-90035394; PubMed-2572549;
RA Munson R.S. Jr., Grass S., Einhorn M., Bailey C., Newell C.;
RT "Comparative analysis of the structures of the outer membrane protein

```

```

RT Pl genes from major clones of Haemophilus influenzae type b.";
RL Infect. Immun. 57:3300-3305(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-7109, AND 6107;
RX MEDLINE-20359342; PubMed-10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC -1- SIMILARITY: STRONG, TO E.COLI LONG-CHAIN FATTY ACID TRANSPORT
CC PROTEIN (PADL).
DR EMBL; M63151; AAA24991.1; -.
DR EMBL; AF260351; AAF97565.1; -.
DR EMBL; AF260350; AAF97564.1; -.
DR PIR; A40183; A40183.
KW Outer membrane; Signal; Transmembrane.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 455 OUTER MEMBRANE PROTEIN PI.
SQ SEQUENCE 455 AA; 49141 MW; 795059130A39D9E1 CRC64;

Query Match          7.3%; Score 13; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 393 GATYKFTPNLSVD 405

RESULT 19
Q9KHI4 PRELIMINARY; PRT; 456 AA.
ID Q9KHI4;
AC Q9KHI4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OMPPI.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID-727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BCH-1;
RX MEDLINE-20359342; PubMed-10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260337; AAF97551.1; -.
SQ SEQUENCE 456 AA; 49402 MW; 8CF3E87CB81DD6E6 CRC64;

Query Match          7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
DB 394 GATYKFTPNLSVD 406

RESULT 20
Q9KHI3 PRELIMINARY; PRT; 456 AA.
ID Q9KHI3;
AC Q9KHI3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DN Outer membrane protein PI.
GN OmpP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260338; AAF97552.1; -.
SQ SEQUENCE 456 AA; 49372 MW; D488A3BA24C8F4EB CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
| | | | | | | | | | | | | | | | | |
DB 394 GATYKFTPNLSVD 406

RESULT 21
Q9KHI2 ID Q9KHI2 PRELIMINARY; PRT; 456 AA.
AC Q9KHI2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OmpP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=3-H-1552;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260339; AAF97553.1; -.
SQ SEQUENCE 456 AA; 49403 MW; F95A977CBCE4F057 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
| | | | | | | | | | | | | | | | | |
DB 394 GATYKFTPNLSVD 406

RESULT 22
Q9KHI1 ID Q9KHI1 PRELIMINARY; PRT; 456 AA.
AC Q9KHI1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Outer membrane protein PI.
```

```
GN OmpP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=17-H-783;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260340; AAF97554.1; -.
SQ SEQUENCE 456 AA; 49308 MW; CECC10E0BB6781A8 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
| | | | | | | | | | | | | | | | | |
DB 394 GATYKFTPNLSVD 406

RESULT 23
Q9KHI0 ID Q9KHI0 PRELIMINARY; PRT; 456 AA.
AC Q9KHI0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OmpP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=21-H-1328;
RC MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260341; AAF97555.1; -.
SQ SEQUENCE 456 AA; 49237 MW; 94A04F0A9A649750 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
| | | | | | | | | | | | | | | | | |
DB 394 GATYKFTPNLSVD 406

RESULT 24
Q9KHH9 ID Q9KHH9 PRELIMINARY; PRT; 456 AA.
AC Q9KHH9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein PI.
GN OmpP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
```

```
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=27-H-1433;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260342; AAF97556.1; -.
SQ SEQUENCE 456 AA; 49239 MW; 35101EFC7A7679D2 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 394 GATYKFTPNLSVD 406

RESULT 25
Q9KHG6 PRELIMINARY; PRT; 456 AA.
AC Q9KHG6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7416;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260347; AAF97561.1; -.
SQ SEQUENCE 456 AA; 49370 MW; 0FA50D60942905D0 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 394 GATYKFTPNLSVD 406

RESULT 26
Q9KHG9 PRELIMINARY; PRT; 456 AA.
AC Q9KHG9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=1071;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260358; AAF97572.1; -.
SQ SEQUENCE 456 AA; 49130 MW; C16D7F3809AC8FA2 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 394 GATYKFTPNLSVD 406

RESULT 27
Q9KHG8 PRELIMINARY; PRT; 456 AA.
AC Q9KHG8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCH-2;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260359; AAF97573.1; -.
SQ SEQUENCE 456 AA; 49075 MW; CF7B3ECE4213C841 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GATYKFTPNLSVD 133
Db 394 GATYKFTPNLSVD 406

RESULT 28
Q9KHG7 PRELIMINARY; PRT; 456 AA.
AC Q9KHG7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein P1.
GN OMPP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO-127;
RX MEDLINE=20359342; PubMed=10899849;
```

RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
 RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
 RT "Variability of outer membrane protein PI and its evaluation as a
 RT vaccine candidate against experimental otitis media due to nontypeable
 RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
 RL Infect. Immun. 68:4505-4517(2000).
 DR EMBL: AF260360; AAF97574.1; -.
 SQ SEQUENCE 456 AA; 49220 MW; 7AC1F56499D94C91 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
 |||||
 Db 394 GATYKFTPNLSVD 406

RESULT 29
 Q9KHG5 PRELIMINARY; PRT; 456 AA.
 AC Q9KHG5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Outer membrane protein PI.
 GN OMPI.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1512A;
 RX MEDLINE=20359342; PubMed=10899849;
 RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
 RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
 RT "Variability of outer membrane protein PI and its evaluation as a
 RT vaccine candidate against experimental otitis media due to nontypeable
 RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
 RL Infect. Immun. 68:4505-4517(2000).
 DR EMBL: AF260362; AAF97576.1; -.
 SQ SEQUENCE 456 AA; 49044 MW; 3581AE61922B511E CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 GATYKFTPNLSVD 133
 |||||
 Db 395 GATYKFTPNLSVD 407

RESULT 30
 Q9KHG4 PRELIMINARY; PRT; 456 AA.
 AC Q9KHG4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Outer membrane protein PI.
 GN OMPI.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=49;
 RX MEDLINE=20359342; PubMed=10899849;
 RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
 RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
 RT "Variability of outer membrane protein PI and its evaluation as a

RT vaccine candidate against experimental otitis media due to nontypeable
 RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
 RL Infect. Immun. 68:4505-4517(2000).
 DR EMBL: AF260363; AAF97577.1; -.
 SQ SEQUENCE 456 AA; 49253 MW; A39BBE8891302566 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
 |||||
 Db 395 GATYKFTPNLSVD 407

RESULT 31
 Q9KHG3 PRELIMINARY; PRT; 456 AA.
 AC Q9KHG3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Outer membrane protein PI.
 GN OMPI.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1-H-1085;
 RX MEDLINE=20359342; PubMed=10899849;
 RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
 RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
 RT "Variability of outer membrane protein PI and its evaluation as a
 RT vaccine candidate against experimental otitis media due to nontypeable
 RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
 RL Infect. Immun. 68:4505-4517(2000).
 DR EMBL: AF260364; AAF97578.1; -.
 SQ SEQUENCE 456 AA; 49133 MW; 0D6AFC5DFEB35014 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
 |||||
 Db 395 GATYKFTPNLSVD 407

RESULT 32
 Q9KHG6 PRELIMINARY; PRT; 457 AA.
 AC Q9KHG6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Outer membrane protein PI.
 GN OMPI.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5;
 RX MEDLINE=20359342; PubMed=10899849;
 RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
 RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
 RT "Variability of outer membrane protein PI and its evaluation as a
 RT vaccine candidate against experimental otitis media due to nontypeable
 RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
 RL Infect. Immun. 68:4505-4517(2000).

DR EMBL; AF260361; AAF97575.1; -;
SQ SEQUENCE 457 AA; 49292 MW; D25CB4D6E81A965F CRC64;

Query Match 7.3%; Score 13; DB 2; Length 457;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||
DB 395 GATYKFTPNLSVD 407

RESULT 33

Q9K397 ID Q9K397 PRELIMINARY; PRT; 458 AA.
AC Q9K397;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT-81, AND 1363;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260356; AAF97570.1; -;
DR EMBL; AF260355; AAF97569.1; -;
SQ SEQUENCE 458 AA; 49674 MW; 1F886C1FAD682F94 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||
DB 396 GATYKFTPNLSVD 408

RESULT 34

Q9K324 ID Q9K324 PRELIMINARY; PRT; 459 AA.
AC Q9K324;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=218, AND 200;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260344; AAF97558.1; -;
DR EMBL; AF260343; AAF97557.1; -;

SQ SEQUENCE 459 AA; 49717 MW; 54490C08E4729870 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||
DB 397 GATYKFTPNLSVD 409

RESULT 35

Q9KHH8 ID Q9KHH8 PRELIMINARY; PRT; 459 AA.
AC Q9KHH8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=667;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260345; AAF97559.1; -;
SQ SEQUENCE 459 AA; 49773 MW; 5E6DCD78F9A8FF8A CRC64;

Query Match 7.3%; Score 13; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||
DB 397 GATYKFTPNLSVD 409

RESULT 36

Q9KHH7 ID Q9KHH7 PRELIMINARY; PRT; 459 AA.
AC Q9KHH7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Outer membrane protein P1.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=199;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RT "Variability of outer membrane protein P1 and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach.";
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260346; AAF97560.1; -;
SQ SEQUENCE 459 AA; 49745 MW; 5A46FF07187E9883 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 459;

Best Local Similarity 100.0%; Pred. No. 7e-05; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||

DB 397 GATYKFTPNLSVD 409

RESULT 37

Q9KHH0 PRELIMINARY; PRT; 460 AA.
AC Q9KHH0
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Long-chain fatty acid transport protein.
GN OMP1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9-H-1194;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach."
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260357; AAF97581.1; -;
SQ SEQUENCE 460 AA; 49719 MW; F1CEBEB597A5CBA7 CRC64;

Query Match 7.3%; Score 13; DB 2; Length 460;

Best Local Similarity 100.0%; Pred. No. 7e-05; Mismatches 0; Indels 0; Gaps 0;

QY 121 GATYKFTPNLSVD 133
|||||

DB 398 GATYKFTPNLSVD 410

RESULT 38

Q9KHF5 PRELIMINARY; PRT; 449 AA.
AC Q9KHF5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Outer membrane protein PI.
GN Omp1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9;
RX MEDLINE=20359342; PubMed=10899849;
RA Bolduc G.R., Bouchet V., Jiang R.Z., Geisselsoder J.,
RA Truong-Bolduc Q.C., Rice P.A., Pelton S.I., Goldstein R.;
RT "Variability of outer membrane protein PI and its evaluation as a
RT vaccine candidate against experimental otitis media due to nontypeable
RT Haemophilus influenzae: an unambiguous, multifaceted approach."
RL Infect. Immun. 68:4505-4517(2000).
DR EMBL; AF260374; AAF97588.1; -;
SQ SEQUENCE 449 AA; 48424 MW; F536A974B0C7F827 CRC64;

Query Match 5.6%; Score 10; DB 2; Length 449;

Best Local Similarity 100.0%; Pred. No. 0.095; Mismatches 0; Indels 0; Gaps 0;

QY 24 ELSGFHQLTD 33
|||||

DB 296 ELSGFHQLTD 305

RESULT 39

Q9KL84 PRELIMINARY; PRT; 412 AA.
AC Q9KL84
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Long-chain fatty acid transport protein.
GN VCA0862.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004414; AAF96760.1; -;
DR TIGR; VCA0862; -;
DR InterPro; IP001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 412 AA; 44784 MW; FB606C929838D891 CRC64;

Query Match 5.0%; Score 9; DB 16; Length 412;

Best Local Similarity 100.0%; Pred. No. 0.98; Mismatches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 WELSGFHQL 31
|||||

DB 261 WELSGFHQL 269

RESULT 40

O28679 PRELIMINARY; PRT; 72 AA.
AC O28679
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Heme D1 biosynthesis protein (NIRD).
GN AFI593.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kachum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

```

RA  Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA  Venter J.C.;
RT  "The complete genome sequence of the hyperthermophilic, sulphate-
RL  reducing archaeon Archaeoglobus fulgidus."
RL  Nature 390:364-370(1997).
DR  EMBL; AE000992; AAB89654.1; -
DR  TIGR; AF1593; -
DR  InterPro; IPR003662; sub_transporter.
DR  PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 72 AA; 8211 MW; 6FBA1C8802970079 CRC64;

Query Match      4.58; Score 8; DB 17; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  171 YGLNLNLYR 178
    |||||
Db   52 YGLNLNLYR 59

Search completed: May 12, 2003, 09:48:49
Job time : 70 secs

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 19, 2003, 18:57:34 ; Search time 1734' Seconds
(without alignments)
3004.269 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 QHNGVLGPYIGKSLTLKP.....ADYTKATANLYGLNLNRYR 179

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlh
-O=/cnp2.1/USPTO_scool/US0995493/runat_12052003_091308_23003/app_query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995493.ecgn_1.1.1687@runat_12052003_091308_23003 -NCPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rdd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	580.5	60.7	1284	1 HIU28808	U28808 Haemophilus
2	571.5	59.7	1350	1 AF260368	AF260368 Haemophil
3	571.5	59.7	1350	1 AF260369	AF260369 Haemophil
4	571.5	59.7	1350	1 AF260370	AF260370 Haemophil
5	570.5	59.6	1347	6 AR085993	AR085993 Sequence
6	570.5	59.6	1350	1 HEAOMP1B	M73494 Haemophilus
7	564	58.9	1356	1 AF260371	AF260371 Haemophil
8	563.5	58.9	1356	1 AF260372	AF260372 Haemophil
9	562.5	58.8	1350	1 AF260374	AF260374 Haemophil
10	561	58.6	1368	1 AF260353	AF260353 Haemophil
11	561	58.6	1368	1 AF260354	AF260354 Haemophil
12	559.5	58.5	1371	1 AF260364	AF260364 Haemophil
13	559	58.4	1365	6 AR085992	AR085992 Sequence
14	559	58.4	1368	1 AF260349	AF260349 Haemophil
15	559	58.4	1368	1 AF260350	AF260350 Haemophil
16	559	58.4	1368	1 AF260351	AF260351 Haemophil
17	559	58.4	1368	1 AF260352	AF260352 Haemophil
18	559	58.4	1368	1 HEAOMP1A	M63151 Haemophilus
19	557	58.2	1368	1 AF260348	AF260348 Haemophil
20	553.5	57.8	1356	1 AF260373	AF260373 Haemophil
21	552.5	57.7	1353	1 AF260375	AF260375 Haemophil
22	552.5	57.7	1353	1 AF260376	AF260376 Haemophil
23	550.5	57.5	1356	1 AF260367	AF260367 Haemophil
24	549.5	57.4	1371	1 AF260358	AF260358 Haemophil
25	549	57.4	1371	1 AF260347	AF260347 Haemophil
26	549	57.4	1371	1 AF260360	AF260360 Haemophil
27	549	57.4	1381	6 AR199524	AR199524 Sequence
28	549	57.4	1560	6 AR085991	AR085991 Sequence
29	549	57.4	1598	1 HEAOMP1	J03381 Haemophilus
30	549	57.4	1598	6 AR089267	AR089267 Sequence
31	549	57.4	1598	6 AR093467	AR093467 Sequence
32	549	57.4	13070	1 U32723	U32723 Haemophilus
33	547	57.2	1356	1 AF260365	AF260365 Haemophil
34	547	57.2	1380	1 AF260343	AF260343 Haemophil
35	547	57.2	1380	1 AF260344	AF260344 Haemophil
36	547	57.2	1380	1 AF260345	AF260345 Haemophil
37	547	57.2	1380	1 AF260346	AF260346 Haemophil
38	545.5	57.0	1371	1 AF260359	AF260359 Haemophil
39	544	56.8	1371	1 AF260337	AF260337 Haemophil
40	544	56.8	1371	1 AF260339	AF260339 Haemophil
41	542.5	56.7	1374	1 AF260361	AF260361 Haemophil
42	539.5	56.4	14766	1 AE006147	AE006147 Pasteurel
43	538	56.2	1371	1 AF260342	AF260342 Haemophil
44	538	56.2	1377	1 AF260355	AF260355 Haemophil
45	538	56.2	1377	1 AF260356	AF260356 Haemophil

ALIGNMENTS

RESULT 1

```

HIU28808
LOCUS      HIU28808      1284 bp      DNA      linear      BCT 10-OCT-1996
DEFINITION Haemophilus influenzae outer membrane protein p1 gene, consensus
sequence, partial cds.
ACCESSION U28808
VERSION   U28808.1 GI:881591
KEYWORDS
SOURCE     Haemophilus influenzae.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
            Haemophilus.
REFERENCE  1 (bases 1 to 1284)
            Reed,R.B., Frost,J.B., Kort,K., Myers,S.D. and Lesse,A.J.
            DNA sequence analysis and restriction fragment length polymorphisms
            of the p1 gene of Haemophilus influenzae biogroup aegyptius
            associated with Brazilian purpuric fever
            Infect. Immun. 64 (9), 3666-3672 (1996)
JOURNAL   96355877
MEDLINE   8751915
PUBMED
REFERENCE  2 (bases 1 to 1284)
            Lesse,A.J.
            Direct Submission
            Submitted (09-JUN-1995) Alan J. Lesse, Infectious Diseases, Buffalo
            VAMC, 151 VAMC, 3495 Bailey Ave, Buffalo, NY 14215, USA
JOURNAL   Location/Qualifiers
FEATURES   Location
            1..1284
            /organism="Haemophilus influenzae"
            /strain="F3029"
            /db_xref="taxon:727"
            /clone="pJBF2B, pJBF3B, and pJBF4C"
            /note="biogroup aegyptius; consensus sequence of three
            clones; pJBF3B and -4C were sequenced in their entirety,
            while pJBF2B was sequenced only in the regions where -3B
            and -4C differed"
            <1..1284
            /codon_start=1
            /transl_table=11
            /product="outer membrane protein p1"
            /protein_id="AAC4459.1"
            /db_xref="GI:881592"
            /translation="AAFLAEVSTSLGRAYAGAAIADNASVYATNPALMSLFKTNQ
            FSVGGYVDSDRMNGDQAAVQNLAKMFTQYGSASERNVVFCAFPNLYFVAPVNDK
            FALGAGMNVFLGKSKYDQADYDIFGGKDLTATNLNSGAYRVTEGLSLGLGVNDK
            YAKAQRNAGLIIVDSKDSQVQNALKGVDPTVPGYLTSDKSVSLQDRANGFG
            WNAGVYNEANRGLAYTHSKVDIDFTDRTATSLKANVKGKGDGLTKLPDYLEL
            SGHLQTDFAVHYSTKYTHSRKLTKLHASFEDGKAFDKELQYNNRSVALGASYNL
            DEKLTLAGIADWDQAASRRHSAAPDPTDRTWYSLGATYKFTPNLSVDLGYAYLKGKK
            VFHKEVQVGLIITANTYTSQAHANLYGLNLNYSF"
            1..18
            /note="forward PCR primer".
primer_bind misc_difference 181
            /note="Probable PCR error, occurs in pJBF3B"
            /replace="t"
misc_difference 411
            /note="Probable PCR error, occurs in pJBF4C"
            /replace="a"
misc_difference 762
            /note="Probable PCR error, occurs in pJBF3B"
            /replace="g"
misc_difference 793
            /note="Probable PCR error, occurs in pJBF3B"
            /replace="a"
misc_difference 1153
            /note="Probable PCR error, occurs in pJBF3B"
            /replace="c"
primer_bind 391 a 220 c 283 g 390 t
            /note="reverse PCR primer"
BASE COUNT 391 a 220 c 283 g 390 t
ORIGIN
Alignment Scores:
Pred. No.:
Score: 1.25e-51 Length: 1284
580.50 Matches: 107

```

```

Percent Similarity: 82.14% Conservative: 31
Best Local Similarity: 63.69% Mismatches: 27
Query Match: 60.66% Indels: 3
DB: 1 Gaps: 2
US-09-995-493-52 (1-179) x HIU28808 (1-1284)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 787 AAGGTGATTTAAACCTTAACCTTAAATGCCAGATTACTTAGAACCTTTACAGTTTCCATCAATTA 846
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 847 ACTGACAAATTTGCTGTGCATTATAGTTATAAATATACCCATTGGAGTCGTTTAACAAAA 906
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 907 TTACATGCTAGCTTCGAAGATGGTAAA-----AAAGCGTTTGATAAAGAATTACAATAC 960
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 961 AGTAATAACTCTCGTTGTCATTAGGGGCAAGTTATATCTTGATGAAAAATTGACCTTA 1020
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1021 CGTGGGGTATTGCTGGGATCAAGCCGCATCTCGTCATCACCGCTAGTGTGCAATTCCA 1080
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1081 GATACCATCCGACCTTGGTATAGTTTATAGGGCAACCTTAATAATTCACGCCGAATTTATCT 1140
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1141 GTTCATCTTGGCTATGCTTACTTAAAGGCAAAAGTTTCACTTTAAAGAAGTACACAA 1200
QY 152 IleLysGlyLeuLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyr 171
Db 1201 GTT---GGCCTCATATAAACAACCGCAATATACACTTCTCAAGCACACGCAATCTTTAC 1257
QY 172 GlyLeuAsnLeuAsnTyrArgPhe 179
Db 1258 GGCTTAACCTTAATATATAGTTTC 1281
RESULT 2
AF260368 1350 bp DNA linear BCT 07-AUG-2000
LOCUS Haemophilus influenzae strain 6255 Outer membrane protein p1
DEFINITION (comp1) gene, complete cds.
ACCESSION AF260368
VERSION AF260368
KEYWORDS AF260368.1 GI:9716628
SOURCE Haemophilus influenzae.
ORGANISM Haemophilus influenzae.
REFERENCE 1 (bases 1 to 1350)
            Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
            Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
            Variability of outer membrane protein p1 and its evaluation as a
            vaccine candidate against experimental otitis media due to
            nontypeable Haemophilus influenzae: an unambiguous, multifaceted
            approach
            Infect. Immun. 68 (8), 4505-4517 (2000)
JOURNAL 20359342
MEDLINE 10899849
PUBMED
REFERENCE 2 (bases 1 to 1350)
            Bouchet,V.
            Direct Submission
            Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
            Finland Laboratory for Infectious Diseases, Boston University
            School of Medicine, Boston Medical Center, 774 Albany Street,
            Boston, MA 02118, USA
JOURNAL Location/Qualifiers
FEATURES

```

[illegible]

44

```

|||||
Db 1321 TACGGCTAAACTTAAATATAGTTTC 1347
RESULT 7
AF260371 1356 bp DNA linear BCT 07-AUG-2000
LOCUS
DEFINITION
Haemophilus influenzae strain 2-H-1038 outer membrane protein p1
(omp1) gene, complete cds.
ACCESSION
AF260371
VERSION
AF260371.1 GI:9716634
KEYWORDS
Haemophilus influenzae.
SOURCE
Haemophilus influenzae
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
REFERENCE
1 (bases 1 to 1356)
AUTHORS
Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
TITLE
Variability of outer membrane protein p1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL
Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE
20359342
PUBMED
10899849
REFERENCE
2 (bases 1 to 1356)
AUTHORS
Bouchet,V.
TITLE
Direct Submission
JOURNAL
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
source
Location/Qualifiers
1..1356
/organism="Haemophilus influenzae"
/strain="2-H-1038"
/serotype="non-typable"
/db_xref="taxon:727"
/country="USA"
1..1356
/gene="omp1"
1..1356
/gene="omp1"
/note="OMP p1"
/codon_start=1
/transl_table=11
/product="outer membrane protein p1"
/protein_id="AAF97585.1"
/db_xref="GI:9716635"
/translation="MKFQSLATAMLLAAGGANAATQLAEVSTSGLGRAVAGBAA
IADNAAVATNPALMSLFKTAQSTGGVYVDSRINNMGDVAASITGSSIKVTNDDVAS
ERNVPCAFVFNLYFVAVNDKLVAGAGMNVNFKLSEYDDSDYDAGVFGKTDLTAIN
LNLGAYRYTEGLSLGLGVNAVYAKAQVERNAGIIADSKVPOAKRALETVPVPTKIP
DYLTSDKSVSLQDRAANGFGWAGVMYQFNEANRIGLAIYHSKVDIDTDRPTATSLG
KGIAGTGNLTFTLPDYLELSCFHLTDKFAVHYSYKYTHWSRLTKLHASEDKG
AFDKLQYSNRSVFLGASTNLDEKUTLRAGIAYDQAASRRHRSRAIPDPTDRTWYSLG
F"
BASE COUNT 416 a 227 c 300 g 412 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7,16e-50 Length: 1356
Score: 564.00 Matches: 108
Percent Similarity: 76.97% Conservative: 29
Best Local Similarity: 60.67% Mismatches: 37
Query Match: 58.93% Indels: 4
DB: 1 Gaps: 3
US-09-995-493-52 (1-179) x AF260371 (1-1356)
Qy 4 GlyValLeuGlyProTyrIleGlyLys--GlySerLeuThrLeuLysLeuProAlaTyr 22
|||||
|||||
|||||

```

```

/gene="ompPl"
/note="OMP Pl"
/codon_start=1
/transl_table=11
/product="outer membrane protein Pl"
/protein_id="AAF97586.1"
/db_xref="GI:9716637"
/transl_table="MKKFNQSLATAMLLAAGGANAFAQLAEVSTSGLRAYAGEAA
IADNSVATNPALMSLEKTAQFSTGCVYVDRINMGVDVSNKATMNTKYGSAQ
ERNVPGAFVFNLYFVAVNDKFAAGMNVNFKLSEYDDYDAGVGGKTDLSAIN
LNLGSKVSVSLQDRAAAGLGNVAVYQFNEANRIGLAYHSHKVDLDFDRTATSLG
DLTSDKSVSLQDRAAAGLGNVAVYQFNEANRIGLAYHSHKVDLDFDRTATSLG
OKSIEAGKTDLTLPDYLELGSFQHLDKFAVHYSYKTHWSRLTKLNANFEDGKA
VFDEKQYNNNSRVALGASYNLEKTLRAGIAYDQAAASHIRSAAPDPTDRTWYSLG
ATYKFTPNLSVDLGYAYLKGKVFKEVQKVGADGLITTNANTSOAHANLYGLNLYS
F"

BASE COUNT      415 a      237 c      295 g      409 t

ORIGIN

Alignment Scores:
Pred. No.:      8.08e-50      Length:      1356
Score:          563.50      Matches:      105
Percent Similarity: 79.64%      Conservative: 28
Best Local Similarity: 62.87%      Mismatches: 31
Query Match:      58.88%      Indels:      3
DB:              1          Gaps:      2

US-09-995-493-52 (1-179) x AF260372 (1-1356)

Qy 14 SerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAsp 33
Db 859 AATTAAACCCCTTACATTTGCCAGATTACTTAGAACCTTCTTGGTTTCCATCAATTAACGCAC 918

Qy 34 GlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArg 53
Db 919 AAATTCGCTGTCATTATAGTTATTAATATACCATTTGGAGTCGTTTAAACAGATTATAC 978

Qy 54 GlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAsp 73
Db 979 GCTAGCTATGAAGATGCTAAG-----AAAGTGTTTCATAAAGAAATTAACAATACAGTAAT 1032

Qy 74 AsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAla 93
Db 1033 AACTCTCGTGTGCTAGGAGGCAAGTTATAATCTTCATGAAATAATTGACCTTACGTCGC 1092

Qy 94 GlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThr 113
Db 1093 GGTATTGCTTACGATCAAGCGCATCTCGTCATCACCGTAGTCGCAATTCACATACC 1152

Qy 114 AspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAsp 133
Db 1153 GATCGCACTTGGTATAGTTAGGGCAACCTATATAATTCACGCCCAATTTATCTGTTGAT 1212

Qy 134 ValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIle--- 152
Db 1213 CTTGGCTATGCTTACTTTAAAGGCAAAAGTTCCATTTAAAGAAAGTACAAAAGTTCGA 1272

Qy 153 LysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGly 172
Db 1273 GATGCTTTGATACAAACCAAGGCAATATACACTTCTCAAGCACACGCAAAATCTTTACGGT 1332

Qy 173 LeuAsnLeuAsnTyrArgPhe 179
Db 1333 TTAACCTAAATATAGTTTC 1353

RESULT 9
AF260374
LOCUS
DEFINITION
Haemophilus influenzae strain 9 outer membrane protein Pl (ompPl)
gene, complete cds.
ACCESSION
AF260374
VERSION
AF260374.1 GI:9716640
KEYWORDS

```

SOURCE

Haemophilus influenzae.
Haemophilus influenzae

Bacteria; Proteobacteria; gamma subdvision; Pasteurellaceae;

REFERENCE

1 (bases 1 to 1350)

AUTHORS

Bolduc, G.R., Bouchet, V., Jiang, R.Z., Gelselsoder, J.,

Truong-Bolduc, O.C., Rice, P.A., Pelton, S.I. and Goldstein, R.

Varliability of outer membrane protein Pl and its evaluation as a vaccine candidate against experimental otitis media due to nontypeable Haemophilus influenzae: an unambiguous, multifaceted approach

Infect. Immun. 68 (8), 4505-4517 (2000)

MEDLINE

20359342

PUBMED

10899849

REFERENCE

2 (bases 1 to 1350)

AUTHORS

Bouchet, V.

TITLE

Direct Submission

JOURNAL

Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell

Finland Laboratory for Infectious Diseases, Boston University

School of Medicine, Boston Medical Center, 774 Albany Street,

Boston, MA 02118, USA

FEATURES

source

gene

1. .1350

/gene="ompPl"

/strain="9"

/serotype="non-typable"

/db_xref="taxon:727"

/country="USA"

1. .1350

/gene="ompPl"

/notes="OMP Pl"

/codon_start=1

/transl_table=11

/product="outer membrane protein Pl"

/protein_id="AAF97586.1"

/db_xref="GI:9716641"

/translation="MKKFNQSLATAMLLAAGGANAFAQLAEVSTSGLRAYAGEAA

IADNSVATNPALMSLEKTAQFSTGCVYVDRINMGVDVSNKATMNTKYGSAQ

RNVPGAFVFNLYFVAVNDKFAAGMNVNFKLSEYDDYDAGVGGKTDLSAINL

NLGSARVTEGLSLGVLGNVAVYQFNEANRIGLAYHSHKVDLDFDRTATSLG

YVTSKDKSGDLTLPDYLELGSFQHLDKFAVHYSYKTHWSRLTKLNANFEDGKA

FDKELQYNNNSRVALGASYNLEKTLRAGIAYDQAAASHIRSAAPDPTDRTWYSLG

TYKFTPNLSVDLGYAYLKGKVFKEVQKVGADGLITTNANTSOAHANLYGLNLYS"

BASE COUNT 417 a 235 c 293 g 405 t

ORIGIN

Alignment Scores:

Pred. No.: 1.02e-49

Score: 562.50

Percent Similarity: 76.70%

Best Local Similarity: 60.80%

Query Match: 58.78%

DB: 1

US-09-995-493-52 (1-179) x AF260374 (1-1350)

Qy 4 GlyValLeuGlyProTyrIleGlySerLeuThrLeuLysLeuProAlaTyrTrp 23

Db 829 GGGGTTATCGGGCGGGG---AAAAAAGGTGATTAAACCCCTTACATTTGCCAGATTACTTA 885

Qy 24 GluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrGlyTyr 43

Db 886 GAACCTTTCTGGTTTCCCATTAATTAACCTTGCAGTGCATTAATAGTATTAATAATAT 945

Qy 44 ThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGlu 63

Db 946 ACCCATTTGGAGTCGTTTAAACAAATTAACCGCTAACCTTCGAGATGGTAAA-----AAA 999

Qy 64 AlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyr 83

Db 1000 GCCTTTGATAAAGATTACATACAGTAATAACTCTCGTGTGCATTAGGGCGCAAGTTAT 1059

Qy 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyAspLysAlaAlaSerLys 103

Db 1060 AATCTTTATGAAAAATTAACTTACCTTACGCTGCGGTATTGCTTACGATCAAGCGCATCTGT 1119

Qy 104 ThrHisLeuSerAlaSerIleProAspThrAspArgMetTyrSerIleGlyAlaThr 123

Db 1120 CATCACCGTAGTCGCAATTCAGATACCGATCGCACTTGGTATAGTTAGGGCAACC 1179

Qy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLys 143

Db 1180 TATAAATTCACCCCTAATTTATGCTGTGCTTACCTTACCTTAAAGGCAACAAA 1239

Qy 144 LysHisPheValGluThrGluAsnIleLysGlyLeuLeuValGluAlaAspTyrThr 163

Db 1240 GTTCACCTTTAAAGAAAGTACAAACAGTAGTGGCCACATAAATAACAAACGCAATTTACACT 1299

Qy 164 ThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179

Db 1300 TCTCAAGCAGCATGCAATCTTTACGGCTTAACCTTAAATTAATAGTTTC 1347

RESULT 10

AF260353

LOCUS AF260353 1368 bp DNA linear BCT 07-AUG-2000

DEFINITION Haemophilus influenzae strain 161 outer membrane protein P1 (ompP1) gene, complete cds.

ACCESSION AF260353

VERSION AF260353.1 GI:9716598

KEYWORDS Haemophilus influenzae.

SOURCE Haemophilus influenzae.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

REFERENCE 1 (bases 1 to 1368)

AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J., Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.

TITLE Variability of outer membrane protein P1 and its evaluation as a vaccine candidate against experimental otitis media due to nontypeable Haemophilus influenzae: an unambiguous, multifaceted approach

JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)

MEDLINE 20359342

PUBMED 10899849

REFERENCE 2 (bases 1 to 1368)

AUTHORS Bouchet,V.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell Finland Laboratory for Infectious Diseases, Boston University School of Medicine, Boston Medical Center, 774 Albany Street, Boston, MA 02118, USA

FEATURES

source Location/Qualifiers

1. .1368

/organism="Haemophilus influenzae"

/strain="161"

/serotype="non-typable"

/db_xref="taxon:727"

/country="Finland"

1. .1368

/gene="ompP1"

1. .1368

/gene="ompP1"

/note="OMP P1"

/codon_start=1

/transl_table=11

/product="outer membrane protein P1"

/protein_id="AAF97567.1"

/db_xref="GI:9716598"

/translat="MKRFNQSIATAMLLAAGGANAAFQIAEVSTSGLGRAVAGEAA

IADNASVYATNPALMSLEKTAQSTGGVYDSRINMGDVTASITATRMNSAKYGSAS

ERNVPGAFVFNLYFVAVNDKFGALGMVNFGLKSEYDDSDYDAGFGKGTDLTAIN

LNLSGAYRVTEGLSLGCVNAVYAKAOVERNAGIANSVNDTQVOTVSLVAPLKL

DQNLPSKDKSVSLQDRAANGFGWAGVYQFNEANRIGLAYHSKVIDIDFTDRATSV

EANVKEGKGNLTLTLPDYLELSGFHQLTDKFAVHYYSKYKTHWSRLTKLHASPEDGK

KAFKLEQYSSNNSVALGASINDELKTLTRAGIAYDQAAASHHRSAAIPDTRTWYSL

KATYKFTPNLSVDLGYAYLKGVKHFKEVKTIGDKRALTLNTTANYTTSQAHANLYGLN

LNYSF"

BASE COUNT 424 a 235 e 293 g 416 t

ORIGIN

Alignment Scores:

Score: 1,49e-49 Length: 1368

Pred. No.: 561.00 Matches: 107

Percent Similarity: 79.65% Conservative: 30

Best Local Similarity: 62.21% Mismatches: 29

Query Match: 58.62% Indels: 6

DB: 1 Gaps: 3

US-09-995-493-52 (1-179) x AF260353 (1-1368)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31

Db 856 AAGGTAATTTAAACCTTACATTCAGATTCAGTACTTGTGGGTTCCATCAATTA 915

Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51

Db 916 ACTGACAAATTTGCAGTCGATATAGTTATAATATACCCATTGGAGTCGTTACAAAA 975

Qy 52 LeuArgGlyLysTyrGluAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71

Db 976 TTACACGCTAGCTTCGAAGATGGTAAG-----AAAGCTTTTGATAAAGAAATTAACAATAC 1029

Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91

Db 1030 AGTAACTCTCGTGTGCATTAGGGCAAGTTATAATCTTGATGAAAAATTAACCTTA 1089

Qy 92 ArgAlaGlyLeuAlaTyAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111

Db 1090 CGTGGCGGTATTGCTTACGATCAAGCTGCATCTCGTCATCACCGTAGTCTCAATTC 1149

Qy 112 AspThrAspArgMetTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131

Db 1150 GATACCGATCGCACTTGGTATAGTTAGTGCAACCTATAAATTCACGCCGAATTTATCT 1209

Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysPheValGluThrGlnAsn 151

Db 1210 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAGTTACCTTTAAAGAGATAAACA 1269

Qy 152 ile-----LysGlyLeuLeuLeu---ValGluAlaAspTyrThrLysAlaThr 167

Db 1270 ATAGTGACAAACGGTCATTGACATTGAATACAACCTGCAAAATATATCTTCAACACAC 1329

Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179

Db 1330 GCAATCTTACGGTTGAATTTAAATTAATATAGTTTC 1365

RESULT 11

AF260354

LOCUS AF260354 1368 bp DNA linear BCT 07-AUG-2000

DEFINITION Haemophilus influenzae strain OT-9 outer membrane protein P1 (ompP1) gene, complete cds.

ACCESSION AF260354

VERSION AF260354.1 GI:9716600

KEYWORDS Haemophilus influenzae.

SOURCE Haemophilus influenzae.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

REFERENCE 1 (bases 1 to 1368)

AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J., Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.

TITLE Variability of outer membrane protein P1 and its evaluation as a vaccine candidate against experimental otitis media due to nontypeable Haemophilus influenzae: an unambiguous, multifaceted approach

JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)

Db 976 TTACACGCTAGCTTCGAGAGTGGTAAG-----AAAGCTTTTGATAAGAAATTACAATAC 1029

Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 :::::::::::::::::::::
 Db 1030 AGTAATAACTCTCGTGTGCATTAGGCGCAAGTTATATCTTGATGAANAATTACCTTA 1089
 Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 :::::::::::::::::::::
 Db 1090 CGTGGGGTATTGCTTACGATCAAGCTGCATCTCGTCATCACCGTAGTGTGCAATTCCTCA 1149
 Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 :::::::::::::::::::::
 Db 1150 GATACCGATCCGACTCTGGTATAGTTAGTGCAACCTATAAATTCACGCCCAATTTATCT 1209
 Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 :::::::::::::::::::::
 Db 1210 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAGTTCTTAAAGAAAGTAAAAACA 1269
 Qy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 :::::::::::::::::::::
 Db 1270 ATAGGTGACAAACGTCATCTGACATTCGAATCAACTGCAAAATTATCTTCTCAAGCACAC 1329
 Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 :::::::::::::::::::::
 Db 1330 GCAAACTCTTTACGGTTTGAATTTAAATTATAGTTTC 1365

RESULT 16
 AF260351 1368 bp DNA linear BCT 07-AUG-2000
 LOCUS
 DEFINITION Haemophilus influenzae strain 7109, outer membrane protein p1
 (omp1) gene, complete cds.
 ACCESSION AF260351
 VERSION AF260351.1 GI:9716594
 KEYWORDS
 SOURCE Haemophilus influenzae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Haemophilus.
 REFERENCE 1 (bases 1 to 1368)
 AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
 Truong-Bolduc,O.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
 TITLE Variability of outer membrane protein p1 and its evaluation as a
 vaccine candidate against experimental otitis media due to
 nontypeable Haemophilus influenzae: an unambiguous, multifaceted
 approach
 JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
 MEDLINE 20359342
 PUBMED 10899849
 REFERENCE 2 (bases 1 to 1368)
 Bouchet,V.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
 Finland Laboratory for Infectious Diseases, Boston University
 School of Medicine, Boston Medical Center, 774 Albany Street,
 Boston, MA 02118, USA
 FEATURES
 source Location/Qualifiers
 1. .1368
 /organism="Haemophilus influenzae"
 /strain="7109"
 /serotype="non-typable"
 /db_xref="taxon:727"
 /country="Sweden"
 1. .1368
 /gene="omp1"
 1. .1368
 /cne="omp1"
 /note="OMP p1"
 /codon_start=1
 /trans_table=11
 /product="outer membrane protein p1"
 /protein_id="AAF97565.1"
 /db_xref="GI:9716595"
 /translation="MKFQNSILATAMLLAAGGANAATFQAEVSTSGLGRAYAGEA

IADNASVYATNPALMSLFTKTAQFSTGGVYVDSRINMNGDVTASLATTKMNSAKYGSAS
 ERNVGCAFPVNLIFYVAPVNDKFPALGAGMNVFGLKSEYDDSDYDAGLEGGKTDLTALN
 LNLSSGVRVTEGLSLGLGVNAVYAKAVERNAGLIANSVNDTQVKTALSVALPLKGL
 DQNLPSKDSKSVSLQDRAAMGFQVAGMVFQNEANRIGLIATHSKVIDLDFDRTATSV
 EANVIREGKKNJUTLPLDYLELSGFGHLDKFAVHSYKYTHMSRLTKLHASEFDYSL
 KATYDKELQYNNRVALGASYNLDEKLTLRAGTAYDQAAASHHRSAAIPDTRTWYSL
 GATYKFTPNLSVDLGYAYLKGKKVHFKEVKTIGDKRTLTLNTTANYTQAHANLYGLN
 LNYSF"

BASE COUNT 427 a 235 c 292 g 414 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,42e-49 Length: 1368
 Score: 559.00 Matches: 106
 Percent Similarity: 78.49% Conservative: 29
 Best Local Similarity: 61.63% Mismatches: 31
 Query Match: 58.41% Indels: 6
 DB: 1 Gaps: 2

US-09-995-493-52 (1-179) x AF260351 (1-1368)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 :::::::::::::::::::::
 Db 856 AAAGTAATTTAACCCCTTACATTCGACAGATTACTTAGAACCTTCTGGGTTCCATCAATTA 915
 Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 :::::::::::::::::::::
 Db 916 ACTCACAAATTTGCGATGTCATTATAGTTATAATATATACCCATTGGAGTCGTTAACAAAA 975
 Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
 :::::::::::::::::::::
 Db 976 TTACACGCTAGCTTCGAGAGTGGTAAG-----AAAGCTTTTGATAAGAAATTACAATAC 1029
 Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 :::::::::::::::::::::
 Db 1030 AGTAATAACTCTCGTGTGTCATTAGGCGCAAGTTATATCTTGATGAAAAATTACCTTA 1089
 Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 :::::::::::::::::::::
 Db 1090 CGTGGGGTATTGCTTACGATCAAGCTGCATCTCGTCATCACCGTAGTGTGCAATTCCTCA 1149
 Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 :::::::::::::::::::::
 Db 1150 GATACCGATCCGACTCTGGTATAGTTAGTGCAACCTATAAATTCACGCCCAATTTATCT 1209
 Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 :::::::::::::::::::::
 Db 1210 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAGTTCATCTTAAAGAAAGTAAAAACA 1269
 Qy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 :::::::::::::::::::::
 Db 1270 ATAGGTGACAAACGTCATCTGACATTCGAATCAACTGCAAAATTATCTTCTCAAGCACAC 1329
 Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 :::::::::::::::::::::
 Db 1330 GCAAACTCTTTACGGTTTGAATTTAAATTATAGTTTC 1365

RESULT 17
 AF260352 1368 bp DNA linear BCT 07-AUG-2000
 LOCUS
 DEFINITION Haemophilus influenzae strain ATCC9795 outer membrane protein p1
 (omp1) gene, complete cds.
 ACCESSION AF260352
 VERSION AF260352.1 GI:9716596
 KEYWORDS
 SOURCE Haemophilus influenzae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Haemophilus.
 REFERENCE 1 (bases 1 to 1368)
 AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
 Truong-Bolduc,O.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
 TITLE Variability of outer membrane protein p1 and its evaluation as a

Qy	12	LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu	31
Db	856	AAAGGTAATTTAAACCTTACATTGCAGATTACTTAGAACCTTCTCGGGTCCATCAATTA	915
Qy	32	ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu	51
Db	916	ACTGACAAATTTGCGATGCTATAGTTATATAATACCCATCGGAGTCGTTTAAACAAA	975
Qy	52	LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr	71
Db	976	TTACACCGCTACCTTCCAGAGTGGTAAG-----AAAGCTTTTGATAAGAAATTAACAATAC	1029
Qy	72	LysAspAsnSerArgPheAlaIleGlyThrTrpTyrSerLeuAsnAlaLeuThrLeu	91
Db	1030	ACTAATAACTCTCGTGTTCATTAGGGCAAGTTATAATCTTGATGAAAAATTAACCTTA	1089
Qy	92	ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro	111
Db	1090	CGTGGGGTATTGCTTACGATCAAGCTGCATCTCGTCATCACCGTAGTCGCAATTC	1149
Qy	112	AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer	131
Db	1150	GATACCGATCGCACTTGGTATAGTTAGTGCAACCTATAATTCACCGCGAATTTATCT	1209
Qy	132	ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn	151
Db	1210	GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAGTTCATCTTAAAGAAAGTAAACAA	1269
Qy	152	IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrLysLysAlaThr	167
Db	1270	ATAGTGACAAAGCGTACATTGACATTTGAATCAACCTGCAAAATTTATCTCTCAAGCAC	1329
Qy	168	AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe	179
Db	1330	GCAATCTTTACGGTTTGAATTTAAATATATAGTTTC	1365
RESULT 19			
AF260348		1368 bp	DNA
LOCUS			linear BCT 07-AUG-2000
DEFINITION			Haemophilus influenzae strain 7004 outer membrane protein p1
ACCESSION			(Omp1) gene, complete cds.
VERSION			AF260348.1
KEYWORDS			GI:9716588
SOURCE			Haemophilus influenzae.
ORGANISM			Haemophilus influenzae
REFERENCE			Bacteria; Proteobacteria; gamma subdivision; Pasteurellales; Haemophilus.
AUTHORS			1 (bases 1 to 1368)
TITLE			Bolduc, G.R., Bouchet, V., Jiang, R.Z., Geisselsoder, J., Truong-Bolduc, Q.C., Rice, P.A., Pelton, S.I. and Goldstein, R. Variability of outer membrane protein p1 and its evaluation as a vaccine candidate against experimental otitis media due to nontypeable Haemophilus influenzae: an unambiguous, multifaceted approach
JOURNAL			Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE			20359342
PUBMED			10899849
REFERENCE			2 (bases 1 to 1368)
AUTHORS			Bouchet, V.
TITLE			Direct Submission
JOURNAL			Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell Finland Laboratory for Infectious Diseases, Boston University School of Medicine, Boston Medical Center, 774 Albany Street, Boston, MA 02118, USA
FEATURES			Location/Qualifiers
source			1. .1368
			/organism="Haemophilus influenzae"
			/strain="7004"
			/serotype="b"
			/db_xref="taxon:727"
			/country="Netherlands"
gene			1. .1368
CDS			/gene="ompP1"
			1. .1368
			/gene="ompP1"
			/note="OMP p1"
			/codon_start=1
			/transl_table=11
			/product="outer membrane protein p1"
			/protein_id="AAF97562.1"
			/db_xref="GI:9716588"
			/translation="MKRFNQSLATAMLLAAGGANAALAEVSTSGLRAYAGEAA
			IADNASVATNPALMSLFKTAQFTGGVYVDSRINMGDVTASLATMNSAKYGSAS
			ERNVPGAEFVNLVYFVNDKFAAGMNVNFGIKSEYDDSYDQKALFGKGLDTAIH
			LNLSGAYRTVEGLSLGCVNAVYAGQVNRNAGIANSVNDQVKTALSVLAPPLKGL
			DONLPSKDKSVVSLQDRAAGFQWAGVYOFNEANRIGLAYHSDKVIDFDTRTATSV
			EANVIKGGKGNLTLPDYLSELSFQHTDKFAVHYSYKYTHSRITKLHASFEDGK
			KAFDKELQYSNNSRVALGSYNLDEKLTLRAGIAYDQAAASHRHSALPTDRTWYSL
			GATKFTPNLSVDLGIYALEGKRVHFKVTIGDKRTLTLNTTNTTYSQAHANDYGLN
			LNYSF"
BASE COUNT	424 a	235 c	293 g 416 t
ORIGIN			
Alignment Scores:			
Pred. No.:	3,92e-49	Length:	1368
Score:	557.00	Matches:	106
Percent Similarity:	77.91%	Conservative:	28
Best Local Similarity:	61.63%	Mismatches:	32
Query Match:	58.20%	Indels:	6
DB:	1	Gaps:	2
US-09-995-493-52 (1-179) x AF260348 (1-1368)			
Qy	12	LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu	31
Db	856	AAAGGTAATTTAAACCTTACATTGCAGATTACTTAGAACCTTCTCGGGTCCATCAATTA	915
Qy	32	ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu	51
Db	916	ACTGACAAATTTGCGATGCTATAGTTATATAATACCCATCGGAGTCGTTTAAACAAA	975
Qy	52	LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr	71
Db	976	TTACACCGCTACCTTCCAGAGTGGTAAG-----AAAGCTTTTGATAAGAAATTAACAATAC	1029
Qy	72	LysAspAsnSerArgPheAlaIleGlyThrTrpTyrSerLeuAsnAlaLeuThrLeu	91
Db	1030	AGTAATAACTCTCGTGTTCATTAGGGCAAGTTATAATCTTGATGAAAAATTAACCTTA	1089
Qy	92	ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro	111
Db	1090	CGTGGGGTATTGCTTACGATCAAGCTGCATCTCGTCATCACCGTAGTCGCAATTC	1149
Qy	112	AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer	131
Db	1150	GATACCGATCGCACTTGGTATAGTTAGTGCAACCTATAATTCACCGCGAATTTATCT	1209
Qy	132	ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn	151
Db	1210	GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAGTTCATCTTAAAGAAAGTAAACAA	1269
Qy	152	IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrLysLysAlaThr	167
Db	1270	ATAGTGACAAAGCGTACATTGACATTTGAATCAACCTGCAAAATTTATCTCTCAAGCAC	1329
Qy	168	AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe	179
Db	1330	GCAATCTTTACGGTTTGAATTTAAATATATAGTTTC	1365
RESULT 20			
AF260373		1356 bp	DNA
LOCUS			linear BCT 07-AUG-2000
DEFINITION			Haemophilus influenzae strain 13-H-1157 outer membrane protein p1
			(ompP1) gene, complete cds.

```

ACCESSION AF260373
VERSION AF260373.1 GI:9716638
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

AF260373
AF260373.1 GI:9716638
Haemophilus influenzae.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
1 (bases 1 to 1356)
Bolduc.G.R., Bouchet.V., Jiang.R.Z., Geisselsoder.J.,
Truong-Bolduc.O.C., Rice,P.A., Pelton,S.I. and Goldstein.R.
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
Infect. Immun. 68 (8), 4505-4517 (2000)
20359342
10899849
2 (bases 1 to 1356)
Bouchet,V.
Direct Submission
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
Location/Qualifiers
1..1356
/organism="Haemophilus influenzae"
/strain="13-H-1157"
/serotype="non-typable"
/db_xref="taxon:727"
/country="USA"
1..1356
/gene="ompP1"
1..1356
/gene="ompP1"
/notes="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/protein_id="AAF97589.1"
/db_xref="GI:9716639"
/translation="MKKNQSLATAMLLAAGGANAAPQLAEVSTSGLRAYAGEAA
IADNVSATNPALMSLEKTAQFSTGGVYDINNMNGVDVAYLQSGSMKFTKYGSAS
ORNVVPGFVFNLYFVAVPNVDFALGAGMNVNFGKSEYDSDYAGIFGGKTDLSAIN
LNLGAYRVTEGLSLGLGVNAVYANAQVNEAGIADSVQDSQVQPALQAVDPQNKAT
GVLTKSKSVLSQDSAAWFGMAGVQVNEANRIGLAYHVKVDIDTDRATSLK
AGAIRAGKGGDITLTPDYLELSGFHOLDKLAVHYSKYTHWSRLTKLNASFEDGKK
AFDKELQVSNNSRVALGASYNLAELTLRAGIAYDQASRHORSRAIPDTRTWSLG
F*
ATYKFTPLNSVDLGLYALGKKVHEKQVKAAGGHIITANTYSQAHANLYGLNLNYS
F*

BASE COUNT 412 a 243 c 295 g 406 t
ORIGIN

Alignment Scores:
Pred. No.: 9,04e-49 Length: 1356
Score: 553.50 Matches: 105
Percent Similarity: 78.11% Conservative: 27
Best Local Similarity: 62.13% Mismatches: 34
Query Match: 57.84% Indels: 3
DB: 1 Gaps: 2

US-09-995-493-52 (1-179), x AF260373 (1-1356)

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGluLeu 31
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
Db 853 AAGGTGATTAAACCTTACATTCGCCAGATTACTTGAACATTTCTGGTTTTCCATCAATTA 912
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 913 ACTGACAACTTCGATGCATATAGTTATATATATACCATTTGGAGTCGCTTTAACAAAA 972
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
||| : : : : : ||||||| : : : : : ||||||| : : : : : |||

```

```

Db 973 TTAACGCTAGCTTCGAAGATGCTAAG-----AAAGCTTTTGTATAAAGAAATTACAAATC 1026
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
: : ||||||| ||||||| : : ||||||| : : ||||||| : : |||||||
Db 1027 AGTAATAACTCTGCTGTCATTAGGGCAAGTTATTAATCTTCTGTAATAATTACCTTA 1086
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
||||| : : ||||||| : : ||||||| : : ||||||| : : |||||||
Db 1087 CGTGGGGTATTCTTACGATCAAGCGGATCTCGTCATCAGCGTAGTGTGCAATTCCTCA 1146
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
||||| : : ||||||| : : ||||||| : : ||||||| : : |||||||
Db 1147 GATACCGATCGCACTTGGTATAGTTAGGTGCAACCTATATAATTCACGCCGAATTTATCT 1206
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsn 151
||||| : : ||||||| : : ||||||| : : ||||||| : : |||||||
Db 1207 GTTCATCTGGCTATGCTTACTTTAAAGGCAAAAAGTTCACCTTTAAAGAGGTACAAAA 1266
QY 152 IleLys---GlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeu 170
||| : : : : : ||||||| : : ||||||| : : ||||||| : : |||||||
Db 1267 GCTGCAAGTGGCCACATATAACAACCGCAATTTACACTTCTCAAGCACACGCAAAATCTT 1326
QY 171 TyrGlyLeuAsnLeuAsnTyrArgPhe 179
||||| ||||||| ||||||| |||
Db 1327 TAGCGCTTAAACTTAAATTTATAGTTTC 1353
RESULT 21
AF260375 AF260375 1353 bp DNA linear BCT 07-AUG-2000
LOCUS Haemophilus influenzae strain 7424 outer membrane protein P1
DEFINITION (OmpP1) gene, complete cds.
ACCESSION AF260375
VERSION AF260375.1 GI:9716642
KEYWORDS Haemophilus influenzae.
SOURCE Haemophilus influenzae
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
REFERENCE 1 (bases 1 to 1353)
AUTHORS Bolduc.G.R., Bouchet.V., Jiang.R.Z., Geisselsoder.J.,
Truong-Bolduc.O.C., Rice,P.A., Pelton,S.I. and Goldstein.R.
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
Infect. Immun. 68 (8), 4505-4517 (2000)
20359342
10899849
2 (bases 1 to 1353)
Bouchet,V.
Direct Submission
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
Location/Qualifiers
1..1353
/organism="Haemophilus influenzae"
/strain="7424"
/serotype="non-typable"
/db_xref="taxon:727"
/country="Kenya"
1..1353
/gene="ompP1"
1..1353
/gene="ompP1"
/notes="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/protein_id="AAF97589.1"
/db_xref="GI:9716643"
/translation="MKKNQSLATAMLLAAGGANAAPQLAEVSTSGLRAYAGEAA

```



```

Db 1141 ATTCCAGATACCGATCGCACTTGGTATAGTTAGGGCGCAACCTATATAATTTACCCCTAAT 1200
QY 130 LeuSerValAlaGlyPheAlaHlsLeuArgGlyLysLysLysHisPheValGluThr 149
Db 1201 TTATCTGTTGATCTTGGCTATGCTTACTTTAAAGGCGCAAAAGTTCACCTTTAAAGAGTA 1260
QY 150 GlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsn 169
Db 1261 AAAACAATAGGTGCCACATATAACAAACGCAAAATACACTTCTCAACGCACATGCAAAAT 1320
QY 170 LeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1321 CTTACGGTGTGAATTTAAATATATAGTTTC 1350

RESULT 23
LOCUS AF260367 1356 bp DNA linear BCT 07-AUG-2000
DEFINITION Haemophilus influenzae strain 4-H-1094 outer membrane protein P1
ACCESSION AF260367
VERSION AF260367.1 GI:9716626
KEYWORDS Haemophilus influenzae.
SOURCE Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
ORGANISM Haemophilus.
REFERENCE 1 (bases 1 to 1356)
AUTHORS Bolduc,G.R., Bouchet,V., Jlang,R.Z., Gelsseisoder,J.,
TITLE Variability of outer membrane protein P1 and its evaluation as a
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE 20359342
PUBMED 10899849
AUTHORS Bouchet,V.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
source
1..1356
Location/Qualifiers
/organism="Haemophilus influenzae"
/strain="4-H-1094"
/serotype="non-typable"
/db_xref="taxon:727"
/country="USA"
1..1356
/gene="ompP1"
1..1356
/gene="ompP1"
/feature="OMP P1"
/transl_table=1
/product="outer membrane protein P1"
/protein_id="AAF97581.1"
/db_xref="GI:9716627"
/translation="MKKEFQSLATAMLLAAGGANAAAFQLAEVSTGLGRAYAGEAA
IADNVATPALMSLPKTAQFTGGYVDSDIRINNGDVKSSTKNLQMLAAKYGSA
ORNVVGFAPVPLNLYFVAVPNKDFALGAGMNVNFGKSEYDSDYAGVGGKTDLSAIN
LNLSCARVTEGLSLGLGVNAVYAKAVERNAGI1IKESVELAKTAFTVAODDEKAIP
YLSKNTSVSLQDRAANGEGNAGVMYQFNEANRIGLAYHSKVIDFDVDRTPLEA
GVIGAKGSDTLTLPDYLELSGFQDLTKLAVHYSKYTHWSRLTKLADPTDRTWSLGA
FDKEIQNSNLSVLAGSYDLYELKTLDRAGIYDQAASHQRHSAIPTDRTWSLGA
P"
BASE COUNT 416 a 230 c 299 g 411 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1..86e-48 Length: 1356
Score: 550.50 Matches: 109
Percent Similarity: 75.84% Conservative: 26
Best local Similarity: 61.24% Mismatches: 38
Query Match: 57.52% Indels: 5
DB: 1 Gaps: 3

US-09-995-493-52 (1-179) x AF260367 (1-1356)
QY 4 GlyValLeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrp 23
Db 829 GGGTTATCGGCGCGCGG---AAAAAAGGTGATTTACCCCTTACATTCGCAGATTACTTA 885
QY 24 GluLeuSerGlyPheHlsGlnLeuThrAspGlnTrpAlaIleHlsTyrSerTyrLysTyr 43
Db 886 GAACCTTCCTGGTTCCATCAATTAACGAACTTGGCGTGCATATAGTTATTAATAT 945
QY 44 ThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGlu 63
Db 946 ACCCATTTGGAGTCGTTTAAACAAATTTAAACGCCAGCTTCGAAGATGCTAAG-----AAA 999
QY 64 AlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyr 83
Db 1000 GCGTTTGAAGAAGATTACAATACAGTAATACTCTCGTTCGATAGGGGCAAGTTAC 1059
QY 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLys 103
Db 1060 GATCTTTATCAAAAATTAACCTTACGTCAGGATTTGCTTACGATCAAGCTGCATCTCGT 1119
QY 104 ThrHlsLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 1120 CATCAGCGTAGTCTGCAATTCAGATACCGATCGCACTTGGTATAGTTTAGGGGAACC 1179
QY 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHlsLeuArgGlyLysLys 143
Db 1180 TATAAATTCACGCCGAATTTATCTGTGATCTTGGCTATGCTTACTTAAAGGCAAAAA 1239
QY 144 LysHisPhe-----ValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAsp 161
Db 1240 GTTCACCTTTAAAGCACTTAAACACAGGTAAACACACTATATTTGAATGCAACTGCAGAT 1299
QY 162 TyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1300 TATATCTCTCAGCACATGCAAAATCTTACGGCTTAAACTTAAATATATAGTTTTC 1353

RESULT 24
LOCUS AF260358 1371 bp DNA linear BCT 07-AUG-2000
DEFINITION Haemophilus influenzae strain 1071 outer membrane protein P1
ACCESSION AF260358
VERSION AF260358.1 GI:9716608
KEYWORDS Haemophilus influenzae.
SOURCE Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
ORGANISM Haemophilus.
REFERENCE 1 (bases 1 to 1371)
AUTHORS Bolduc,G.R., Bouchet,V., Jlang,R.Z., Gelsseisoder,J.,
TITLE Variability of outer membrane protein P1 and its evaluation as a
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE 20359342
PUBMED 10899849
AUTHORS Bouchet,V.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University

```

School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
Location/Qualifiers
1. .1371
/organism="Haemophilus influenzae"
/strain="1071"
/serotype="non-typable"
/db_xref="taxon:727"
/country="Finland"
1. .1371
/gene="ompP1"
1. .1371
/gene="ompP1"
/note="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/protein_id="AAF97572.1"
/db_xref="GI:9716609"
/translation="MKFNQSLATAMLLAAGGANAFAQLAEVSTSGLGRAVAGEAA
IADNASVATNPALMSLFKTAQFSTGGVYVDSRINMGDVASVVIISNMVKATQDQSA
SARNVPMWCFTKSLRCARNDFALCAGMNVFGLKSEYDDSYDAGIFGKGLDLSAI
NLNSGAYRYTEGLSLGLGVNAVYAKAQVERNAGTSDSVDNSQLTGVLSTLAKPKE
LNKHLPSKDSVSLQDRAANGFAGMAGVYQNEANRIGLAYHSKVDIDFTDRATS
LEAGVTCAGTGTTLTPDYLSESGFQHLTKLAVHYSYKTYTHWSRLTKLNASFEDG
KKAFDELQYSNNSRVALGASYNLYEKLTLRAGTAYDQAAASRHSRAAIPDPTDRTWYS
LGATYKFTPNLSVDLGAYLKGKKVHFKEGTIADKSGSLTSTANTYTSQAHANLYGL
NLNYSF"

BASE COUNT 411 a 241 c 300 g 419 t
ORIGIN

Alignment Scores:
Pred. No.: 2,4e-48 Length: 1371
Score: 549.50 Matches: 111
Percent Similarity: 76.80% Conservativeness: 28
Best Local Similarity: 61.33% Mismatches: 33
Query Match: 57.42% Indels: 9
DB: 1 Gaps: 5

US-09-995-493-52 (1-179) x AFB260358 (1-1371)

Qy 4 GlyValLeuGlyProTyrIleGlyLys---GlySerLeuThrLeuLysLeuProAlaTyr 22
Db 838 GGGGTTATCGGG-----CGGGGAAACAGGTAAATTAACCCCTACATGCCAGATTAC 891

Qy 23 TrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTyrAlaIleHisTyrSerTyrLys 42
Db 892 TTAGAACCTTCTGGTTCCATCAATTAACGCACAACTTGGCGTGCATTATAGTTATATAA 951

Qy 43 TyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyr 62
Db 952 TATACCCATTGGAGTCGTTTAAACAAATTTACATGCCAGCTTCGAAGATGGTAAAG----- 1005

Qy 63 GluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThr 82
Db 1006 AAACGGTTTGATAAGAAATACATACAGTAACTCTCTGTTCATAGGGGCGAAGT 1065

Qy 83 TyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSer 102
Db 1066 TATAATCTTTATGAAAAATTAACCTTACGTGCGGTATTGCTTACGATCAAGCTGCATCT 1125

Qy 103 LysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAla 122
Db 1126 CGTCATCAGCGTACTGTGCAATTTCCAGATACCCATCGCACTTCGTATAGTTTAGGGCA 1185

Qy 123 ThrTyrLysPheThrProAspLeuSerValAspValGlyPheAlaHisLeuArgGlyLys 142
Db 1186 ACCTATAAATTTACGCCGAATTTATCTGTGATGCTTGGCTATGCTTACCTAAAGGCCAA 1245

Qy 143 LysLysHisPheValGluThrGlnAsnIle-----LysGlyLeuLeuValGlu 159
Db 1246 AAAGTTCACTTTAAAGGAAGGAAAAACAAGTACGACAAATCTGCATTGCCATTGCACTTCA 1305

Qy 160 ---AlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArg 178
Db 1306 ACTGCAAAATACACTTCTCAAGCACACGCAAAATCTTTACGGCTTAAACTTAAATATAGT 1365

Qy 179 Phe 179
Db 1366 TTC 1368

RESULT 25
AF260347
LOCUS
DEFINITION
Haemophilus influenzae strain 7416 outer membrane protein P1
ACCESSION
AF260347
VERSION
AF260347.1 GI:9716586
KEYWORDS
Haemophilus influenzae.
SOURCE
Haemophilus influenzae
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE
1 (bases 1 to 1371)
AUTHORS
Boliduc, G.R., Bouchet, V., Jiang, R.Z., Geisselsoder, J.,
Truong-Bolduc, Q.C., Rice, P.A., Pelton, S.I., and Goldstein, R.
TITLE
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL
Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE
20359342
PUBMED
10899849
REFERENCE
2 (bases 1 to 1371)
AUTHORS
Bouchet, V.
TITLE
Direct Submission
JOURNAL
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
Location/Qualifiers
1. .1371
/organism="Haemophilus influenzae"
/strain="7416"
/serotype="a"
/db_xref="taxon:727"
/country="Kenya"
1. .1371
/gene="ompP1"
1. .1371
/gene="ompP1"
/note="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/protein_id="AAF97561.1"
/db_xref="GI:9716587"
/translation="MKFNQSLATAMLLAAGGANAFAQLAEVSTSGLGRAVAGEAA
IADNASVATNPALMSLFKTAQFSTGGVYIDSRINMGDVDSAAIISNMKATKDGSA
SERNVPGAEPNLYFVAPYDNGFALCAGMNVFGLKSEYDDSYDAGVFGKGLDLSAI
NLNSGAYRYTEGLSLGLGVNAVYAKAQVERNAGIADSVKDNQVKTALTVQEPKLF
LDKLYPSKDSVSLQDRAANGFAGMAGVYQNEANRIGLAYHSKVDIDFTDRATS
VEANVTKAGKGGDLTLTPDYLSESGFQHLTKLAVHYSYKTYHWSRLTKLNASFEDG
KKAFDELQYSNNSRVALGASYNLYEKLTLRAGTAYDQAAASRHSRAAIPDPTDRTWYS
LGATYKFTPNLSVDLGAYLKGKKVHFKEGTIADKSGSLTSTANTYTSQAHANLYGL
NLNYSF"

BASE COUNT 426 a 230 c 292 g 423 t
ORIGIN

Alignment Scores:
Pred. No.: 2,71e-48 Length: 1371
Score: 549.00 Matches: 106
Percent Similarity: 77.91% Conservativeness: 28
Best Local Similarity: 61.63% Mismatches: 32
Query Match: 57.37% Indels: 6
DB: 1 Gaps: 3

```

US-09-995-493-52 (1-179) x AF260347 (1-1371)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
DB 859 AAAGTGATTTAAACCTTACATTGCCAGATACCTTGAACCTTCTGGTTCCATCAATTA 918
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
DB 919 ACTGACAAACTTCTGCTGCATATAGTATATAATATACCATTTGGAGTCTGTTAAACAAA 978
QY 52 LeuArgLysLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
DB 979 TTAACCGCTAGCTTCCAGATGTAA-----AAAGCTTTTGAATAAGATTTACAATAC 1032
QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAlaLeuThrLeu 91
DB 1033 AGTAATAACTCTCGTGTACATTAGGGCAACCTATAAATTCACGGCAATTTATCT 1212
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaLeuSerLysThrHisLeuSerAlaSerIlePro 131
DB 1093 GATACCGGATTCGCTTACATTAGGGCAACCTATAAATTCACGGCAATTTATCT 1212
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
DB 1213 GTTGATCTTGGCTATGCTTAAAGGCAACCAAAAGATTTACCTTTAAAGAAAGTAACTATCT 1212
QY 152 Ile-----LysGlyLeu---LeuLeuValGluAlaAspTyrThrLysAlaThr 167
DB 1273 ATAGGTAAGAATCTCGCTTGGCTTGGCACTTACAACTGCAAAATTTACACTTCTCAAGCACAC 1332
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
DB 1333 GCAAATCTTTACGGCTTAAACTTAAATTTATAGTTTC 1368

RESULT 26
AF260360
LOCUS
DEFINITION
Haemophilus influenzae strain LO-127 outer membrane protein P1
ACCESSION
AF260360
VERSION
AF260360.1 GI:9716612
KEYWORDS
Haemophilus influenzae.
SOURCE
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE
1 (bases 1 to 1371)
BOLDUC,G.R., BOUCHET,V., JIANG,R.Z., GELSSLSODER,J.,
TRUONG-BOLDUC,Q.C., RICE,P.A., PELTON,S.I. and GOLDSTEIN,R.
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE
20359342
PUBMED
10899849
REFERENCE
2 (bases 1 to 1371)
Bouchet,V.
Direct Submission
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA -02118, USA
LOCATION/Qualifiers
1. 1371
/organism="Haemophilus influenzae"
/strain="LO-127"
/serotype="non-typable"

FEATURES
source
/db_xref="taxon:727"
/country="Canada"
1. 1371
/genes="ompP1"
1. 1371
/genes="ompP1"
/notes="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/protein_id="AAF97574.1"
/db_xref="GI:9716613"
/translation="MKKFNQSLTATMLAAGGANAFAQLAEVSTSGIGRAYAGBA
IADNASVAVATNPALMSLFKTAQSTGGVYVDSTINNGNVPSSVTNTSMKTTMDGSA
SSRNVPAGAFVFNLYFVAPNDKFKALGAGMNVNFKLSEYDDSDAGVFGKTDLSAI
NNLGLSKDKSVSLQDRAAGFGNVMYQFNEANRIGLAYHSKYDLDFTDRTATS
VZANVTKAGKGGDLTLTLPDLELGSFHQTLKLAHYSTYKTHRSAAIPDFTWYS
KKAFTKELQYSNNRVALGASYNLEKLTLAGIAYDQAAHQHRSAAIPDFTWYS
LGATYKFTPNLSVDLGYAYLKGKHFKEVNTIGRESGLPLTTTANYTTSQAHANLYGL
NLNYSF"
BASE COUNT 422 a 244 c 290 g 414 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 2,71e-48 Length: 1371
Score: 549.00 Matches: 107
Percent Similarity: 76.74% Conservative: 25
Best Local Similarity: 62.21% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: 1 Gaps: 3
US-09-995-493-52 (1-179) x AF260360 (1-1371)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
DB 859 AAAGTGATTTAAACCTTACATTGCCAGATACCTTGAACCTTCTGGTTCCATCAATTA 918
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
DB 919 ACTGACAAACTTCCGCTGCATATAGTATATAATATACCATTTGGAGTCTGTTAAACAAA 978
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
DB 979 TTAACCGCTAGCTTCCAGATGTAA-----AAAGCTTTTGAATAAGATTTACAATAC 1032
QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAlaLeuThrLeu 91
DB 1033 AGTAATAACTCTCGTGTACATTAGGGCAACCTATAAATTCACGGCAATTTATCT 1212
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaLeuSerLysThrHisLeuSerAlaSerIlePro 111
DB 1093 CGTGGCGGATTCGCTTACCATCAAGCGGCATCTCGTCATCAGCGTAGTGTGCAATTTCCA 1152
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysLysPheThrProAsnLeuSer 131
DB 1153 GATACCGCGCATCTGGTATAGTTAGGGCAACCTATAAATTCACGGCAATTTATCT 1212
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
DB 1213 GTTGATCTTGGCTATGCTTAAAGGCAACCAAAAGATTTACCTTTAAAGAAAGTAACTATCT 1212
QY 152 Ile-----LysGlyLeu---LeuLeuValGluAlaAspTyrThrLysAlaThr 167
DB 1273 ATAGGTAAGAATCTCGCTTGGCTTGGCACTTACAACTGCAAAATTTACACTTCTCAAGCACAC 1332
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
DB 1333 GCAAATCTTTACGGCTTAAACTTAAATTTATAGTTTC 1368

RESULT 27
AR199524
LOCUS
AR199524 1381 bp DNA linear PAT 20-APR-2002

```

DEFINITION Sequence 5 from patent US 6355450.
ACCESSION AR199524
VERSION AR199524.1 GI:20249598
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1381)
AUTHORS Fleischmann, R.D., Adams, M.D., White, O., Smith, H.O. and Venter, J. Craig.
TITLE Computer readable genomic sequence of Haemophilus influenzae Rd. fragments thereof, and uses thereof
JOURNAL Patent: US 6355450-A 5 12-MAR-2002;
FEATURES Location/Qualifiers
1. 1381
BASE COUNT 424 a 237 c 297 g 423 t
ORIGIN
Alignment Scores:
Pred. No.: 2,74e-48 Length: 1381
Score: 549.00 Matches: 106
Percent Similarity: 77.91% Conservative: 28
Best Local Similarity: 61.63% Mismatches: 32
Query Match: 57.37% Indels: 6
DB: Gaps: 3
US-09-995-493-52 (1-179) x AR199524 (1-1381)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 872 AAAGGTGATTTAAACCTTACATTCGCAGATTACTTAGAACCTTTCTGGTTTCCCAATTA 931
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 932 ACTGACAACTGCTGCTGCATTATAGTTATAAATATACCATTCGAGTCGTTTACAAAA 991
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 992 TTAACCGCTAGCTTCCGAAGATGGTAAA-----AAAGCTTTTGATAAAGAATTACAATAC 1045
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1046 AGTAATAACTCTCGTGGTTGCATTAGGGCAAGTTATAATCTTGATGAAAAATTGACCTTA 1105
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1106 CGTGGGCTATTGCTTACGATCAGCTGCATCTCGTCATCAGCTAGTGTCTCAATTTCA 1165
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1166 GATACCGATCGCACCTGGTATAGTTAGGGCAACCTATAAATTCACGCCGAATTTATCT 1225
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1226 GTTGATCTTGCTATGCTTACTTAAAGCCAAAAGTTCACTTTAAAGAGATAAAACA 1285
QY 152 Ile-----LysGlyLeuLeuLeu----ValGluAlaAspTyrThrThrLysAlaThr 167
Db 1286 ATAGTGACGAACTTCACTTACATTGAATACAACTGCAAAATATATACTTCTCAACGACAT 1345
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1346 GCAAACTTTACGGTTTAAACTTAAATTAATATAGTTTC 1381
RESULT 28
AR085991
LOCUS AR085991 1560 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5985288.
ACCESSION AR085991
VERSION AR085991.1 GI:10012757
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Munson, R.S. Jr., Grass, S., Chong, P., Yang, Y.-P., Fahim, R., Charles, Sja, D. Yan., McVerry, P. and Klein, M.
TITLE Outer membrane protein p1 and peptides of Haemophilus influenzae type B
JOURNAL Patent: US 5985288-A 1 16-NOV-1999;
FEATURES Location/Qualifiers
1. 1560
BASE COUNT 487 a 271 c 321 g 481 t
ORIGIN
Alignment Scores:
Pred. No.: 3,15e-48 Length: 1560
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservative: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: Gaps: 2
US-09-995-493-52 (1-179) x AR085991 (1-1560)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 922 AAAGGTAATTTAAACCTTACATTCGCAGATTACTTAGAACCTTTCTGGTTTCCCAATTA 981
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 982 ACTGACAACTTGCAGTCAGTTATAGTTATAAATATACCATTCGAGTCGTTTACAAAA 1041
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 1042 TTACATGCCAGCTTCGAAGATGGTAAA-----AAAGCTTTTGATAAAGAATTACAATAC 1095
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1096 AGTAATAACTCTCGTGGTTGCATTAGGGCAAGTTATAATCTTGATGAAAAATTGACCTTA 1155
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1156 CGTGGGCTATTGCTTACATCAGCGGCATCTCGTCATCAGCTAGTGTCTCAATTTCA 1215
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1216 GATACCGATCGCACCTGGTATAGTTAGTGCAACCTATAAATTCACGCCGAATTTATCT 1275
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1276 GTTGATCTTGCTATGCTTACTTAAAGCCAAAAGTTCACTTTAAAGAGATAAAACA 1335
QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1336 ATAGTGACAACTGATACATTGACATTGAATACAACTGCAAAATATATACTTCTCAACGACAC 1395
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1396 GCAAACTTTACGGTTTGAATTTAAATTAATATAGTTTC 1431
RESULT 29
HEAOMPP1
LOCUS HEAOMPP1 1598 bp DNA linear BCT 26-APR-1993
DEFINITION Haemophilus influenzae type b, outer membrane protein (omp) p1 gene, complete cds.
ACCESSION J03381
VERSION J03381.1 GI:148952
KEYWORDS Outer membrane protein.
SOURCE H. influenzae type b DNA, clone pKSM188.
ORGANISM Haemophilus influenzae
REFERENCE 1 (bases 1 to 1598)

AUTHORS Munson,R. Jr. and Grass,S.
 TITLE Purification, cloning, and sequence of outer membrane protein P1 of
 Haemophilus influenzae type b
 JOURNAL Infect. Immun. 56 (9), 2235-2242 (1988)
 MEDLINE 88314258
 PUBMED 2842261
 COMMENT Submitted in computer readable form by R.Munson 01-JUL-1988.
 FEATURES Location/Qualifiers

source
 1. .1598

/organism="Haemophilus influenzae"

/db_xref="taxon:727"

72. .1451

/note="outer membrane protein p1 precursor"

/codon_start=1

/transl_table=11

/protein_id="AAA24990.1"

/db_xref="GI:148953"

/translation="MKKFNQSLATAMLLAAGGANAAPQAEVSTGSLGRAYAGEAA
 IADNASVATNPALMSLFTKFTAGVGVYIDSRINMGVDVTSYAQITNQIGMKATKD
 GASQRNVVPGAFVNLVFPVNDKFGALGMNVNFGKSEYDSDYVFGGKTDL
 SAINLNGAYRVTEGLSLGLGVNAVYAKAQVYRNAGLIADSKVNOITSALSTQOEP
 FDLKKYLPSSKDKSVVSLQDRAAGFGNAGVYOFNEANRIGLAVHKSVDIDFADRT
 ATSLKANVKEGCKGNLFTLPDYLELSEFHLQTLKLVHYSYKTYHWSRLTKLASP
 ECKAFADKELQTSNNSRVALGASINLEKTLRAGIAYDQAASRHRSAALPDDTDT
 WYSLGATYKFTPLNSVDLGYAYLKGKVKFEVKTIGDKRTLTLNTANTYTSQAHANL
 YGLNLNYSF"

72. .137

/note="outer membrane protein p1 signal peptide"

138. .1448

/product="outer membrane protein p1"

BASE COUNT 500 a 278 c 329 g 491 t

ORIGIN 122 bp upstream of PstI site.

Alignment Scores:

Pred. No.: 3,24e-48 Length: 1598
 Score: 549.00 Matches: 105
 Percent Similarity: 76.74% Conservative: 27
 Best Local Similarity: 61.05% Mismatches: 34
 Query Match: 57.37% Indels: 6
 DB: 1 Gaps: 2

US-09-995-493-52 (1-179) x HEAOMP1 (1-1598)

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 939 AAAGGTAATTAACTTTACATTCGCCAGATTACTTAGAACCTTCTGGTTCCATCAATTA 998
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 Db 999 ACTGACAACTTCGAGTGCATTTAGTATAATATATACCCATTGGAGTCGTTTAAACAAA 1058
 QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
 Db 1059 TTACATGCCAGCTTCGAAGATGTTAA-----AAAGCTTTTGATAAAGAAATTACAATAC 1112
 QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 Db 1113 AGTAATAACTCTCGTTCGATAGGGCAAGTTATAATCTTTATGAAAAATTGACCTTA 1172
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 Db 1173 CGTGGCGGATTCGTTACGATCAAGCGGCATCTCGTCATCACCGTAGTCGCAATTTCCA 1232
 QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 Db 1233 GATACCGATCGCACTCGGTATAGTTAGGTGCAACCTTATAAATTCACGCCGAATTTATCT 1292
 QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 Db 1293 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAAGTTCACCTTAAAGAAAGTAAAAACA 1352
 QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrLysAlaThr 167
 Db 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179

Db 1353 ATAGTGCACAAACGTACATTGACATTGAATACAACTGCAAAATTACTTCTCAAGCACAC 1412
 QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1413 GCAAATCTTTACGGTTTGAATTAAATATATAGTTTC 1448
 RESULT 30
 AR089267
 LOCUS AR089267 1598 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 26 from patent US 5994066.
 ACCESSION AR089267
 VERSION AR089267.1 GI:10016024
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1598)
 AUTHORS Bergeron,M.G., Picard,F.J., Ouellette,M. and ROY,P.H.
 TITLE Species-specific and universal DNA probes and amplification primers
 to rapidly detect and identify common bacterial pathogens and
 associated antibiotic resistance genes from clinical specimens for
 routine diagnosis in microbiology laboratories
 JOURNAL Patent: US 5994066-A 26 30-NOV-1999;
 FEATURES Location/Qualifiers
 source 1. .1598
 BASE COUNT 500 a 278 c 329 g 491 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,24e-48 Length: 1598
 Score: 549.00 Matches: 105
 Percent Similarity: 76.74% Conservative: 27
 Best Local Similarity: 61.05% Mismatches: 34
 Query Match: 57.37% Indels: 6
 DB: 1 Gaps: 2
 US-09-995-493-52 (1-179) x AR089267 (1-1598)
 QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 939 AAAGGTAATTAACTTTACATTCGCCAGATTACTTAGAACCTTCTGGTTCCATCAATTA 998
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 Db 999 ACTGACAACTTCGAGTGCATTTAGTATAATATATACCCATTGGAGTCGTTTAAACAAA 1058
 QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
 Db 1059 TTACATGCCAGCTTCGAAGATGTTAA-----AAAGCTTTTGATAAAGAAATTACAATAC 1112
 QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 Db 1113 AGTAATAACTCTCGTTCGATAGGGCAAGTTATAATCTTTATGAAAAATTGACCTTA 1172
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 Db 1173 CGTGGCGGATTCGTTACGATCAAGCGGCATCTCGTCATCACCGTAGTCGCAATTTCCA 1232
 QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 Db 1233 GATACCGATCGCACTCGGTATAGTTAGGTGCAACCTTATAAATTCACGCCGAATTTATCT 1292
 QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 Db 1293 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAAGTTCACCTTAAAGAAAGTAAAAACA 1352
 QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrLysAlaThr 167
 Db 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179

Db 1413 CCAATCTTTACGGTTTGAATTTAAATATAGTTTC 1448
RESULT 31
AR093467
LOCUS AR093467 1598 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 26 from patent US 6001564.
ACCESSION AR093467
VERSION AR093467.1 GI:10020216
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1598)
Unclassified.
Bergeon,M.G., Ouellette,M. and Roy,P.H.
TITLE Specific and universal DNA probes and amplification primers to rapidly detect and identify common bacterial pathogens and associated antibiotic resistance genes from clinical specimens for routine diagnosis in microbiology laboratories
JOURNAL patent: US 6001564-A 26 14-DEC-1999;
FEATURES
Location/Qualifiers
1..1598
/organism="unknown"
BASE COUNT 500 a 278 c 329 g 491 t
ORIGIN
Alignment Scores:
Pred. No.: 3,24e-48 Length: 1598
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservative: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
Gaps: 2
Db:
US-09-995-493-52 (1-179) x AR093467 (1-1598)
Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisLeu 31
Db 939 AAGGTAAATTAACTTTACATTCGCCAGATTACTTAGAACCTTCCTGGTTCCATCAATTA 998
Qy 32 ThrAspGlyTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 999 ACTGACAAACTCGAGTCGATATAGTTATAATATATACCCATGGCGTGTAAACAAA 1058
Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 1059 TTACATGCCAGCTTCGAAGATGGTAAA-----AAAGCTTTTGATAAGAAATTAACAATAC 1112
Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1113 AGTAATACTCTCGTTGCAATTAGGGCAAGTTATATCTTTATGAAAATTTGACCTTA 1172
Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1173 CGTGGGTATTGCTTAGCATCAAGCGCATCTCGTCATCATCCGCTAGTGTGCAATTCCA 1232
Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1233 GATACCGATCGCACTTGGTATAGTTAGTGCACACTATAAATTCACGCCGAATTTATCT 1292
Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1293 GTTGATCTTGCTATGCTTACTTAAAGGCCAAAGAAAGTTCACTTTAAGAGAGTAAACA 1352
Qy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1353 ATAGGTGACAAACGTACATTCACATTTGAATCAACTGCAATATTATCTCTCAGCACAC 1412
Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1413 GCAAAATCTTTACGGTTTGAATTTAAATATAGTTTC 1448
RESULT 32
U32723

LOCUS 13070 bp DNA linear BCT 29-MAY-1998
DEFINITION Haemophilus Influenzae Rd section 38 of 163 of the complete genome.
ACCESSION U32723 L42023
VERSION U32723.1 GI:1573363
KEYWORDS
SOURCE Haemophilus Influenzae Rd.
ORGANISM Haemophilus Influenzae Rd
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.
REFERENCE
1 (bases 1 to 13070)
Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Merrick,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A., Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A., Gacey,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodek,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E., Cotton,M.D., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrman,J.V., Smith,H.O., Smith,H.O. and Venter,J.C.
TITLE Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL Science 269 (5223), 496-512 (1995)
MEDLINE 95350630
PUBMED 7542800
REFERENCE
2 (bases 1 to 13070)
Tatsovsky,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Borodovsky,M., Rudd,K.E. and Koonin,E.V.
TITLE Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
JOURNAL Curr. Biol. 6 (3), 279-291 (1996)
MEDLINE 96398784
PUBMED 8805245
REFERENCE
3 (bases 1 to 13070)
White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
4 (bases 1 to 13070)
White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatsovsky et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
5 (bases 1 to 13070)
White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D., Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start
COMMENT On Sep 30, 1996 this sequence version replaced gi:1221066.
FEATURES
Location/Qualifiers
1..13070
/organism="Haemophilus Influenzae Rd"
/db_xref="taxon:71421"
complement(275..1012)
/gene="HI0392"
complement(275..1012)
/gene="HI0392"
/note="similar to GP:1518853 percent identity: 31.46; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="O-antigen acetylase, putative"
/protein_id="AAC2051.1"
/db_xref="GI:1573364"
/translation="MDIFFVISGLITGIITLQONSFSLSQYTRRIKRIYPAFIT VMALVSFIASAIYFNDFNFKURKTIELAIFLSNFIYGLTQGYFDLSANPNVLHWS

```

LAVEGQYLIIPPLIILILAYKKREVKVLFIITLILFFILLATSFVSANFYKEVLHQPNI
IYVLSNLPPELVGSLALAIYHNSKNKQSVNNILAILSTLLFLSCLFLMNNIA
FIPGIFTLIPCIFTALIIHTTSQNNIR"
complement(1195..2286)
/ gene= "HI0393"
CDS
complement(1195..2286)
/ note= "HI0393"
PID:416043 similar to SP:P31216 GB:U02423 GB:X61941 PID:311421
similarity: 81.54; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical GTP-binding protein"
/protein_id="AAC22052.1"
/db_xref="GI:1573365"
/transl_table=11
/translation="MGFKGIVGLPNVGKSTLFNALTRAGIEAANYPFCTIETPNTGVV
PMPRLDALAIEVPERILPTMEFVDIAGIVAGASGEGIGNKFANIRETDAIGH
VYRCFENDDIVHAGKIDPLDDIDITINTELALADLSDSCERAQIRKQKRAKGDKEAKF
ELUSWEKILPVLENAGVMKSVGDKDEELQAIKSYNFLTUKPTMYTANWNEDGFENPNY
LDRVREIAAKEAGAVVPCAAIESEIAELDDKEKVEFFQDLGIEEPGLNRVIRGAYAL
LNQIYFTAGVREAVMTVSGATAPKAAAVIHTDFEKFIRAEVIAYEDFIQFNGEN
GAKGAKWRLEGGKDYIVODGVMHFRFNV"
complement(2335..2919)
/ gene= "HI0394"
CDS
complement(2335..2919)
/ note= "HI0394"
PID:165197 similar to SP:P23932 GB:X61941 PID:581202 GB:U00096
similarity: 64.21; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="peptidyl-trna hydrolase (pth)"
/protein_id="AAC22053.1"
/db_xref="GI:1573366"
/transl_table=11
/translation="MSEIKLIVGLNPGDKYTDTRHAGEWLIERLARFENVSLAPES
KFGFTARTLVNGKEVRLVPTTFNLSCKAVCALASFYRIKPEELVILVDELDPGP
TARKLQGGGGHGLKDIYVQGLNNFNRYLRIGIGHGPHRDVLVAGYVLRNKPSPADR
DALEKVLDEATDCVEMIFRDGMVKATNRLNSPKI"
3086..3394
/ gene= "HI0395"
CDS
3086..3394
/ note= "HI0395"
PID:1033114 similar to SP:P52119 GB:U00096
similarity: 55.32; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22054.1"
/db_xref="GI:1573377"
/transl_table=11
/translation="MNQINIEIAYAPPERYLYKSFQVDEGTVQTAITQSGLISOPFE
IDLSTNKGIFSRPTKLTDLVLEKGDRIEYRPLADPKETIRKRAAEQAQAKDEKGA
"
3397..4611
/ gene= "HI0396"
CDS
3397..4611
/ note= "HI0396"
PID:1788971 similar to SP:P27431 GB:U00096 PID:1787373 percent
identity: 32.58; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22055.1"
/db_xref="GI:1573367"
/transl_table=11
/translation="MTALSSVDFCLPEHTPEIFLRDYWKQKPLVIRNGLPEIVGOFPE
PDQIIELAQNEVDVARTLVKDDMKVFFSLSEKDFQKLPKWSVLVONLEQWSPDE
LAQLWNKFGFIQWQKRDIMVSYAPKGVSGKHGYDEYDFLVQGYGHRRMQVQKWCDA
SFYKFPKQIRIFDQMGDLNIDEVMPNGDILYIPARMAHYGVAEDDCLTFSPGLRYPN
LSTNLIDGISKGFCHODPLNLFEDPLRLSOSQORTGLADENIQAOKQLLDKLALH
SEAFDTLQAVASAVSSRRYELLYSDCMCDPDEVRISILEEDGAFQSQNNCKLLYTE
NPIRYANGEWELDLNIIIESEVLKRLSDGESLDWAFSLDLANKTDPETSMDLLLSDI
CNWVDDGVALIE"

```

```

gene
CDS
4613..5932
/ gene= "HI0397"
4613..5932
/ note= "HI0397"
PID:1788856 similar to GB:J02599 SP:P04994 PID:148275 GB:U00096
similarity: 57.69; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease VII, large subunit (xseA)"
/protein_id="AAC22056.1"
/db_xref="GI:1573368"
/transl_table=11
/translation="MSDNIYSVSQLNSAARQMLEGNFCQIWLITGEISNTQPVSQGHY
LGLKDENAQVRCAMFRMKNLRFVAPQNGQVILVRANVSLYTPRGDYQIITIDSMHPAG
EFLAQQQFALKMLAAEGFLQNLKTLPHFSKAVGIITSTGAALQDILHILARRD
PSLKVVYPTAVQGEATAEIIVQMIELANARQEVDLIVRGGSUDELWCNEEVEA
RAIFSTPLIISAVGHETDVTIADFVADLRAPTSPSAAAELVSRNODLQLQRLHQOOR
LMAFDRLETRKSRQLKQLALRLOHQHPONLRAQAQKNEQLTHRLQLAILAQFENTQ
QKPTALSVRKQNPPIRYRIQYQQGLEUKLRNFCNVQTERQNKLATLCGKLDGL
SPKLKARGYSIAENPQGAIVSVKDVNGQDFITTTQVADGKIVSKVL"
6109..6762
/ gene= "HI0398"
6109..6762
/ note= "HI0398"
PID:882564 similar to SP:P36651 PID:882564 GB:U00096
similarity: 54.00; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22057.1"
/db_xref="GI:1573369"
/transl_table=11
/translation="MQFWRKQMSIEQHFSSQDIEILGEQTLVEGFPTLKRIOFKKHLF
AGQSGVTVRELLIKGASAVIAYDPKDSVILVEQVRIGAAVHPESHSPWLELLIA
GAVEKEPEDVALRESEEGIQVKNLTHCLSVNPSGSAVERIHLFAGEVDSADAK
GTHGLAEENEDIKVHVVKREQAYQVMCEKIDNGIAVIGLQHLQNLNYAQLOQSWKRS"
6772..7596
/ gene= "HI0399"
6772..7596
/ note= "HI0399"
PID:882562 similar to GB:D16557 SP:P36650 PID:453396
similarity: 53.26; identified
by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="lacZ expression regulator (icc)"
/protein_id="AAC22058.1"
/db_xref="GI:1573370"
/transl_table=11
/translation="MKNTFYQAEKPVIKLLQITDPHLEKDESAELLGVNTQASPAQV
LKEIQENNEFDVILATGDLVQDSDSEGYIRFVEMMKPNKPVFIPTGNHDFQPKMAE
FLNQPMNAKHLGLGEHWQALLDSQVYQPHGQLSQHLDLLKTKETLGNKRPRTLV
VLHHLLPTNSAWLDOHNLNLSHELAEVLAPFTNVKAILYGHITHQVNSMNGYQVMA
TPATCICQPKDCQYFESLDTLPQWREIELHSDGSIRTEVKRIQQAEEFFPMQEEGI"
complement(7694..7981)

```

Alignment Scores:

Pred. No.:	3-76e-47	Length:	13070
Score:	549.00	Matches:	106
Percent Similarity:	77.91%	Conservative:	28
Best Local Similarity:	61.63%	Mismatches:	32
Query Match:	57.3%	Indels:	6
DB:	1	Gaps:	3

US-09-995-493-52 (1-179) x U32723 (1-13070)

Qy	12	LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu	31
Db	9055	AAAGGTATTAAACCCCTTACATTCGCCAGATTACTTGAACCTTTCTGGTTTTCCATCAATTA	9114
Qy	32	ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu	51
Db	9115	ACTGACAAACTTCTGTCGCACTATAGTTATATAATATATACCATTCGGAGTCGTTTAAACAAA	9174

TITLE
 Truong-Bolduc, O.C., Rice, P.A., Pelton, S.I. and Goldstein, R.
 Variability of outer membrane protein p1 and its evaluation as a
 vaccine candidate against experimental otitis media due to
 nontypeable *Haemophilus influenzae*: an unambiguous, multifaceted
 approach

JOURNAL
 Infect. Immun. 68 (8), 4505-4517 (2000)

MEDLINE
 20359342

PUBMED
 10899849

REFERENCE
 2 (bases 1 to 1380)

AUTHORS
 Bouchet, V.

TITLE
 Direct Submission

JOURNAL
 Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
 Finland Laboratory for Infectious Diseases, Boston University
 School of Medicine, Boston Medical Center, 774 Albany Street,
 Boston, MA 02118, USA

FEATURES
 Location/Qualifiers

source

1..1380

/organism="Haemophilus influenzae"

/strain="200"

/serotype="non-typable"

/db_xref="taxon:727"

/country="Finland"

1..1380

/gene="ompP1"

1..1380

/gene="ompP1"

/note="OMP p1"

/codon_start=1

/transl_table=11

/product="outer membrane protein p1"

/protein_id="AAF97557.1"

/db_xref="GI:9716579"

/translation="MKFNQSLIATAMLLAAGGANAFAQLAEVSTSGLRVAGAA

IADNSWATNPALMSLEKTAQSTGCVYIDSRNNMGDVNSDAITTSSTSGVKTIKH

GSASQRNVIPGAPVNPVYFVAPVNDKFAAGAGNVNFKLSEYDDSDAGIEGKTDL

SATNLNSGAYRTEGLSLGLGNVAYAKAQRNAGIIVDSLKDNQVQTALSVQKEP

LKYLHKYLPKSDTSVSLQDRAGWFGNAGVYQFNEANRIGLATHSVKVIDFAORT

ATSVKADVIKAGKKGDLTLTPDYLELSGFHQLTKLAVHYSYKTHWRLTKLNASF

EDGKKADEKLOYNNRSIALGASYNLYEKLTRAGIAYDQASRRHQRSAIIPDTRT

WYSLGATYKFTPNLSVDLGYAYLKGKKVHKVKEVKTIGDERSIALNTANTYTSQAHNL

YGLNLNYSF"

BASE COUNT 430 a 236 c 293 g 421 t

ORIGIN

Alignment Scores:

Pred. No.: 4,43e-48 Length: 1380

Score: 547.00 Matches: 106

Percent Similarity: 77.33% Conservative: 27

Best Local Similarity: 61.63% Mismatches: 33

Query Match: 57.16% Indels: 6

DB: 1 Gaps: 3

US-09-995-493-52 (1-179) x AF260343 (1-1380)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31

Db 868 AAAGTGATTAAACCCCTACATTGCCAGATTACTTAGAACCTTCTGCTTCCATCAATTA 927

Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrTrpGluAlaPheHisGlnLeu 51

Db 928 ACTGACAACTTCGGTGCATATAGTTATAATATACCCATGCGATGCTTTTAAACAAA 987

Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluLys 71

Db 988 TTAACCGCTAGCTTCGAAGATGGTAAA-----AAAGCGTTTGATGAAGAATTACAAATAC 1041

Qy 72 LysAspHisSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91

Db 1042 AGTAATAACTCTCCTATTGCTATTAGGGCAAGTTATATCTTTATGAAAAACTCACCTTA 1101

Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111

Db 1102 CGTCGGGGTATTGCTTACGATCAAGCGGCATCTCGTCATCAGCGCTAGTGGCTGCAATTTCCA 1161

Oy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 Db 1162 GATACCATCGCATCTGGTATAGTTTAGGGCAACCTATTAATTCACGCCCAATTTATCT 1221
 Oy 132 ValaspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 Db 1222 GTTGATCTTGGCTATGCTTAAAGGCAAAAAGTTCACTTTAAAGAAATAAAAACA 1281
 Oy 152 Ile-----LysGlyLeuLeuLeu--ValGluAlaAspTyrThrThrLysAlaThr 167
 Db 1282 ATAGTGACGACACCTTCACTTGCATTGAATACAACTCAATATATCTTCAAGCACAT 1341
 Oy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1342 GCAATCTTTACGGCTTAACCTTAATATATAGTTTC 1377

RESULT 35
 AF260344
 LOCUS
 DEFINITION
 Haemophilus influenzae strain 218 outer membrane protein p1 (ompP1)
 gene, complete cds.
 ACCESSION
 AF260344
 VERSION
 AF260344.1
 KEYWORDS
 GI:9716580
 SOURCE
 Haemophilus influenzae.
 ORGANISM
 Haemophilus influenzae.
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Haemophilus.

REFERENCE
 1 (bases 1 to 1380)
 AUTHORS
 Bolduc, G.R., Bouchet, V., Jiang, R.Z., Geisselsoder, J.,
 Truong-Bolduc, O.C., Rice, P.A., Pelton, S.I. and Goldstein, R.
 TITLE
 Variability of outer membrane protein p1 and its evaluation as a
 vaccine candidate against experimental otitis media due to
 nontypeable *Haemophilus influenzae*: an unambiguous, multifaceted
 approach
 JOURNAL
 Infect. Immun. 68 (8), 4505-4517 (2000)
 MEDLINE
 20359342
 PUBMED
 10899849
 REFERENCE
 2 (bases 1 to 1380)
 AUTHORS
 Bouchet, V.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
 Finland Laboratory for Infectious Diseases, Boston University
 School of Medicine, Boston Medical Center, 774 Albany Street,
 Boston, MA 02118, USA

FEATURES
 source
 1..1380
 /organism="Haemophilus influenzae"
 /strain="218"
 /serotype="non-typable"
 /db_xref="taxon:727"
 /country="Finland"
 1..1380
 /gene="ompP1"
 1..1380
 /gene="ompP1"
 /note="OMP p1"
 /codon_start=1
 /transl_table=11
 /product="outer membrane protein p1"
 /protein_id="AAF97558.1"
 /db_xref="GI:9716581"

TRANSLATION="MKFNQSLIATAMLLAAGGANAFAQLAEVSTSGLRVAGAA
 IADNSWATNPALMSLEKTAQSTGCVYIDSRNNMGDVNSDAITTSSTSGVKTIKH
 GSASQRNVIPGAPVNPVYFVAPVNDKFAAGAGNVNFKLSEYDDSDAGIEGKTDL
 SATNLNSGAYRTEGLSLGLGNVAYAKAQRNAGIIVDSLKDNQVQTALSVQKEP
 LKYLHKYLPKSDTSVSLQDRAGWFGNAGVYQFNEANRIGLATHSVKVIDFAORT
 ATSVKADVIKAGKKGDLTLTPDYLELSGFHQLTKLAVHYSYKTHWRLTKLNASF
 EDGKKADEKLOYNNRSIALGASYNLYEKLTRAGIAYDQASRRHQRSAIIPDTRT
 WYSLGATYKFTPNLSVDLGYAYLKGKKVHKVKEVKTIGDERSIALNTANTYTSQAHNL
 YGLNLNYSF"

BASE COUNT 430 a 236 c 293 g 421 t

gene
 CDS
 1..1380
 /gene="ompP1"
 1..1380
 /gene="ompP1"
 /note="OMP p1"
 /codon_start=1
 /transl_table=11
 /product="outer membrane protein p1"
 /protein_id="AAF97558.1"
 /db_xref="GI:9716581"
 /translation="MKFNQSLIATAMLLAAGGANAFAQLAEVSTSGLRVAGAA
 IADNSWATNPALMSLEKTAQSTGCVYIDSRNNMGDVNSDAITTSSTSGVKTIKH
 GSASQRNVIPGAPVNPVYFVAPVNDKFAAGAGNVNFKLSEYDDSDAGIEGKTDL
 SATNLNSGAYRTEGLSLGLGNVAYAKAQRNAGIIVDSLKDNQVQTALSVQKEP
 LKYLHKYLPKSDTSVSLQDRAGWFGNAGVYQFNEANRIGLATHSVKVIDFAORT
 ATSVKADVIKAGKKGDLTLTPDYLELSGFHQLTKLAVHYSYKTHWRLTKLNASF
 EDGKKADEKLOYNNRSIALGASYNLYEKLTRAGIAYDQASRRHQRSAIIPDTRT
 WYSLGATYKFTPNLSVDLGYAYLKGKKVHKVKEVKTIGDERSIALNTANTYTSQAHNL
 YGLNLNYSF"

BASE COUNT 430 a 236 c 293 g 421 t

ORIGIN

Alignment Scores:
Pred. No.: 4,43e-48 Length: 1380
Score: 547.00 Matches: 106
Percent Similarity: 77.33% Conservative: 27
Best Local Similarity: 61.63% Mismatches: 33
Query Match: 57.16% Indels: 6
DB: 1 Gaps: 3

US-09-995-493-52 (1-179) x AF260344 (1-1380)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 868 AAGGTGATTAAACCTTACATTCAGCAGATTAAGTATGAGTTCCTGCTTCCATCAATTA 927

Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTyrSerArgPheLysGlu 51
Db 928 ACTGACAACTTCGCGTGCATTAAGTATTAATATACCCATTGAGTTCGTTTACAAAA 987

Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
Db 988 TAAACGCTAGCTTCGAAGATGGTAAA-----AAAGCGTTTGATAAGAAATTACAATAC 1041

Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1042 AGTAATAACTCTCGTATTGCATTAGGGCAAGTTATAATCTTTATGAAAAACTCACCTTA 1101

Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1102 CGTGGGGTATTGCTTAGCATCAACGCGCATCTCGTCATCAGCTAGTGTGCAATTCAC 1161

Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1162 GATACCGATCGCAGCTGGTATAGTTAGGGCAACCTATAAATTCACGCCGAATTTATCT 1221

Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1222 GTTGATCTGGCTATGCTTACTTAAAGGCAAAAAGTTCACTTTAAAGAAAGTAAACA 1281

Qy 152 Ile-----LysGlyLeuLeuLeu---ValGluAlaAspTyrThrThrLysAlaThr 167
Db 1282 ATAGTGACGACGAGCTTCCATTGCAATTAACAACCTGCAAAATTTACTTCTCAACACAT 1341

Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1342 GCAAATCTTAGCGCTTAACCTTAAATATATAGTTTC 1377

RESULT 36
AF260345 LOCUS AF260345 1380 bp DNA linear BCT 07-AUG-2000
DEFINITION Haemophilus influenzae strain 667 outer membrane protein PI (ompP1) gene, complete cds.
ACCESSION AF260345
VERSION AF260345.1 GI:9716582
KEYWORDS Haemophilus influenzae.
SOURCE Haemophilus influenzae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellales; Haemophilus.
REFERENCE 1 (bases 1 to 1380)
AUTHORS Boldus,G.R., Bouchet,V., Jlang,R.Z., Geisselsoder,J., Truong-Bolduc,Q.C., Rhee,P.A., Pelton,S.I. and Goldstein,R.
TITLE Variability of outer membrane protein PI and its evaluation as a vaccine candidate against experimental otitis media due to nontypeable Haemophilus influenzae: an unambiguous, multifaceted approach
JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE 20359342
PUBMED 10899849
REFERENCE 2 (bases 1 to 1380)
AUTHORS Bouchet,V.
TITLE Direct Submission

JOURNAL

Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA

FEATURES

Source 1..1380 Location/Qualifiers
/organism="Haemophilus influenzae"
/strain="667"
/serotype="non-typable"
/db_xref="taxon:727"
/country="Finland"
1..1380
/gene="ompP1"
1..1380
/gene="ompP1"
/note="OMP p1"
/codon_start=1
/transl_table=11
/product="outer membrane protein p1"
/protein_id="AAF97559.1"
/db_xref="GI:9716583"
/translation="MKFPNQSLATAMLLAVGGANAAAFQALAEVSTSGIGRAYAGEAA
IADNASVATNPALMSLPKTAQFSTGGVYIDSRINMGDVSNDATITSTNSGVKTIKH
GSASORWIPGAFVFNLYFVAVNDKFAAGMNVNFKLSEYDSDYDAGIFGSKTDL
SAINLUSGATRVTEGSLGUVNAVYAKAQRNAGLIVDSKDNQYQTALSVOKER
LKYLRKYLPSKDTSVSLQDRAAMGFGNAGVMYQFNEANRIGLAYHSKVDIDFADRT
ATSVADYIKAGKGDLTLTLPDYLELSCFQLDKLAVHYSYKTHWSRITKLNASF
EDGKATFDEQLQYNNRSIALGSYLNLYKLTLAGIAYDQAAASHQSSAIPDTRT
WYSLGATYKFTPNLSVDLGYAYLKGKKVHKVEKVTIGDERSIALNTTANYTSQAHNL
YGLNLYSF"

BASE COUNT 431 a 235 c 292 g 422 t
ORIGIN

Alignment Scores:

Pred. No.: 4,43e-48 Length: 1380
Score: 547.00 Matches: 106
Percent Similarity: 77.33% Conservative: 27
Best Local Similarity: 61.63% Mismatches: 33
Query Match: 57.16% Indels: 6
DB: 1 Gaps: 3

US-09-995-493-52 (1-179) x AF260345 (1-1380)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 868 AAGGTGATTAAACCTTACATTCAGCAGATTAAGTATGAGTTCCTGCTTCCATCAATTA 927

Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTyrSerArgPheLysGlu 51
Db 928 ACTGACAACTTCGCGTGCATTAAGTATTAATATACCCATTGAGTTCGTTTACAAAA 987

Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
Db 988 TTAACGCTAGCTTCGAAGATGGTAAA-----AAAGCGTTTGATAAGAAATTACAATAC 1041

Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1042 AGTAATAACTCTCGTATTGCATTAGGGCAAGTTATAATCTTTATGAAAAACTCACCTTA 1101

Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1102 CGTGGGGTATTGCTTAGCATCAACGCGCATCTCGTCATCAGCTAGTGTGCAATTCAC 1161

Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1162 GATACCGATCGCAGCTGGTATAGTTAGGGCAACCTATAAATTCACGCCGAATTTATCT 1221

Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1222 GTTGATCTGGCTATGCTTACTTAAAGGCAAAAAGTTCACTTTAAAGAAAGTAAACA 1281

Qy 152 Ile-----LysGlyLeuLeuLeu---ValGluAlaAspTyrThrThrLysAlaThr 167
Db 1282 ATAGTGACGACGAGCTTCCATTGCAATTAACAACCTGCAAAATTTACTTCTCAACACAT 1341

```

Db 1282 ATAGGTGACGACGTTCACTTCGATTGAATACAACTGCAAAATATATACTTCTCAAGCACAT 1341
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179 linear BCT 07-AUG-2000
||||| 1380 bp DNA outer membrane protein P1 (ompP1)
Db 1342 GCAAACTCTTACGGCTTAAACTTAAATATATAGTTTC 1377

RESULT 37
AF260346 1380 bp DNA linear BCT 07-AUG-2000
LOCUS AF260346 gene, complete cds.
DEFINITION Haemophilus influenzae strain 199 outer membrane protein P1 (ompP1)
ACCESSION AF260346
VERSION AF260346.1 GI:9716584
SOURCE Haemophilus influenzae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE 1 (bases 1 to 1380)
AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
TITLE Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE 20359342
PUBMED 10899849
REFERENCE 2 (bases 1 to 1380)
AUTHORS Bouchet,V.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
source
1..1380
Location/Qualifiers
/organism="Haemophilus influenzae"
/strain="199"
/serotype="non-typable"
/db_xref="taxon:727"
/country="Finland"
1..1380
/gene="ompP1"
1..1380
/gene="ompP1"
/note="OMP P1"
/codon_start=1
/transl_table=11
/product="outer membrane protein P1"
/db_xref="GI:9716585"
/translat="MKKFNOSLLATAMLLAAGGANAAAFQAEVSTSGLGRAYAGEAA
IADNASVATPALMSLEKTAQFTGGVYIDSRINMGDVSADALITTSNGVKTIKH
GSASORNVIPGAFVFNLYFVAPNDKFAAGMNVNFGKSEYDDSDYDAGIFGGKTDL
SAINLUSGAVRTVEGLSLGIVANAYAKAGVERNAGIIVPSLDKNQVOTDLSVOKEP
LKYLHKYLPKSTSVYSLQDRAAGFGWAGVMYQFNEANRIGLAYHKSVDLDFADRT
ATSVADYAIKAGKGDLTLTPDYLELSGFHQLTDKLVAVHSYKYTHWSRLTKLNASF
EDGKAFDKELQYNNRSRIALGASVLYEKLTRAGIAYDQAASHQBSAIPDTRT
WYSLGATYKFTPNLSVDGLYALYKGGKVFKEVKTIGDERSLALNTYTSQAHANL
YGLNLNYSF"
BASE COUNT 432 a 237 c 291 g 420 t
ORIGIN
Alignment Scores:
Pred. No.: 4.43e-48 Length: 1380
Score: 547.00 Matches: 106
Percent Similarity: 77.33% Conservative: 27
Best Local Similarity: 61.63% Mismatches: 33
Query Match: 57.16% Indels: 6
DB: 1 Gaps: 3
US-09-995-493-52 (1-179) x AF260346 (1-1380)

```

```

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 868 AAAGGTGATTTAAACCTTACATTGCCAGATTTACTTAAAGATTTCTGGTTCCATCAATTA 927
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 928 ACTGACAAACTTTCGGTGCATTATAGTTATAAATATACCATTTGGAGTCGTTTAAACAAA 987
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 988 TTAAACGCTAGCTTCGAAGATGTTAAA-----AAAGCGCTTTGATAAAGAAATTTACATA 1041
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAlaLeuLeuLeu 91
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1042 AGTAATAACTTCGTATTCATTAGGGCAAGTTATATCTTTATGAAAACCTCACCTTA 1101
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaIleSerLysThrHisLeuSerAlaSerIlePro 111
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1102 CGTGGCGGTATTCGTACGATCAAGCGCATCTCGTCATCAGCGTAGTCGTCGAATTC 1161
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1162 GATACCGATCGCACTTGGTATAGTTTAGGGCAACCTATAAATTCACGCCGAATTTATCT 1221
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysPheValGluThrGlnAsn 151
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1222 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAAGATTCACCTTTAAAGAGTAAAAACA 1281
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 152 Ile-----LysGlyLeuLeuLeu-----ValGluAlaAspTyrThrLysAlaThr 167
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1282 ATAGGTGACGAAGCTTGCATTGCAATGCAACATGCAAAATATATCTTCTCAAGCACAT 1341
||||| 1371 bp DNA linear BCT 07-AUG-2000
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179 linear BCT 07-AUG-2000
||||| 1371 bp DNA linear BCT 07-AUG-2000
Db 1342 GCAAACTCTTACGGCTTAAACTTAAATATATAGTTTC 1377
||||| 1371 bp DNA linear BCT 07-AUG-2000
RESULT 38
AF260359 1371 bp DNA linear BCT 07-AUG-2000
LOCUS AF260359 Haemophilus influenzae strain BCH-2 outer membrane protein P1
DEFINITION (ompP1) gene, complete cds.
ACCESSION AF260359
VERSION AF260359.1 GI:9716610
KEYWORDS Haemophilus influenzae.
SOURCE Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
ORGANISM Haemophilus.
REFERENCE 1 (bases 1 to 1371)
AUTHORS Bolduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
TITLE Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
Variability of outer membrane protein P1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
JOURNAL Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE 20359342
PUBMED 10899849
REFERENCE 2 (bases 1 to 1371)
AUTHORS Bouchet,V.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
source
1..1371
Location/Qualifiers
/organism="Haemophilus influenzae"
/strain="BCH-2"
/serotype="non-typable"
/db_xref="taxon:727"
/country="USA"

```



```

Db 922 GACAACTTGCAGTCGATTATAGTTATAAATATATACCATTTGGAGTCGTTTAAACAAATTA 981
Qy 53 ArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLys 72
Db 982 AACGCTAGCTCGAAGCTGAAG-----AAAGCGTTTGATAAAGAAATTACAATACAGT 1035
Qy 73 AspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArg 92
Db 1036 AATAACTTCGCTGTCATAGGGCAAGCTATAACCTTGATGAAAAATTGACCTTACGTT 1095
Qy 93 AlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHlsLeuSerAlaSerIleProAsp 112
Db 1096 GCGGGTATTGCTTACCATCAAGCGCATCTCGTCATCACGCTGCTGCAATTCAGAT 1155
Qy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
Db 1156 ACCGATCGCACCTGGTATAGTTAGGGCAACCTATAAATTCACCCCTAATTTATCTGTT 1215
Qy 133 AspValGlyPheAlaHlsLeuArgGlyLysLysLysHlsPheValGluThrGlnAsnIle 152
Db 1216 GATCTTGGCTATGCTTACTTAAAGGCAAAAGTTTCACTTAAAGAAAGTAAACAAATA 1275
Qy 153 -----LysGlyLeuLeuLeu-----ValGluAlaAspTyrThrThrLysAlaThrAla 168
Db 1276 GGTGCAAAACGTTTCATCTGCATTTAGTACAACTGCAATTTACTTCTCAAGCACACGCA 1335
Qy 169 AsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1336 AATCTTTACGGTTTGAATTTAAATTTATAGTTTC 1368

RESULT 40
AF260339
LOCUS
DEFINITION
Haemophilus influenzae strain 3-H-1552 outer membrane protein p1
ACCESSION
AF260339
VERSION
AF260339.1 GI:9716570
KEYWORDS
Haemophilus influenzae.
SOURCE
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE
1 (bases 1 to 1371)
Boiduc,G.R., Bouchet,V., Jiang,R.Z., Geisselsoder,J.,
Truong-Bolduc,Q.C., Rice,P.A., Pelton,S.I. and Goldstein,R.
Variability of outer membrane protein p1 and its evaluation as a
vaccine candidate against experimental otitis media due to
nontypeable Haemophilus influenzae: an unambiguous, multifaceted
approach
Infect. Immun. 68 (8), 4505-4517 (2000)
MEDLINE
20359342
PUBMED
10899849
REFERENCE
2 (bases 1 to 1371)
Bouchet,V.
Direct Submission
Submitted (25-APR-2000) Section of Molecular Genetics, The Maxwell
Finland Laboratory for Infectious Diseases, Boston University
School of Medicine, Boston Medical Center, 774 Albany Street,
Boston, MA 02118, USA
FEATURES
source
1..1371
/organism="Haemophilus influenzae"
/strain="3-H-1552"
/serotype="non-typable"
/db_xref="taxon:727"
/country="USA"
1..1371
/gene="omp1"
1..1371
/gene="omp1"
/notes="OMP p1"
/codon_start=1
/transl_table=11

```

```

/product="outer membrane protein p1"
/protein_id="AAF97553.1"
/db_xref="GI:9716571"
/translation="MKKFNOSLLATAMLLAAGNAAAFOLAEVSTVSGLGRAYAGEAA
IADNASVYATNPALMSLFTKQFVGVYVDSRINMGDGVDSYALLSDSIKYTNDGSA
SARNVVPAPVNLFFVAPVNDKFAAGMNVNFKLSEYDDSYDAGVFGGKTDLTAI
NLNLSGAYRVTEGLSGLVNAVYAKAQVERNAGLIADTVKDNQYKVDLTVQOEPLKY
INKYLPKSDTSVSLQDRAAGFGNAGVMYQFNEANRIGLAYHKSVDLIDFTDRTATS
LEANYINASKTGDLTLPDYLELSGPHQLDKLAVHYSYKYTHSRVLTKLNASPEDG
KKAPDKELQYSNNSRVALGASYNLDEKILTLAGIAYDOAASPHORSAIPLDTRTWYS
LGATYKPTPNLSVDLGAYALGKKVHFKEVTIGDKRSLALSTTANTYTSQAIANLYGL
NLNVSF"
BASE COUNT 424 a 233 c 289 g 425 t
ORIGIN
Alignment Scores:
Pred. No.: 9,08e-48 Length: 1371
Score: 544.00 Matches: 105
Percent Similarity: 77.78% Conservative: 28
Best Local Similarity: 61.40% Mismatches: 32
Query Match: 56.84% Indels: 6
DB: 1 Gaps: 3
US-09-995-493-52 (1-179) x AF260339 (1-1371)
Qy 13 GlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHlsGlnLeuThr 32
Db 862 GGTGATTTAACCTTACATTGCCAGATTACTTAGAATCTTCTGGTTCCATCAATTA 921
Qy 33 AspGlnTrpAlaIleHlsTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeu 52
Db 922 GACAACTTGCAGTCGATTATAGTTATAAATATACCATTTGCAGTCGTTTAAACAAATTA 981
Qy 53 ArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLys 72
Db 982 AACGCTAGCTTTCGAAGATGGTAAG-----AAAGCGTTTGTGATAAGAAATTACAATACAGT 1035
Qy 73 AspAsnSerArgPheAlaIleGlyThrTyrThrSerLeuAsnAspAlaLeuThrLeuArg 92
Db 1036 AATAACTCTCGTTCGATTTAGGGCAAGCTTATAACCTTGATGAAAAATTGACCTTACGT 1095
Qy 93 AlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHlsLeuSerAlaSerIleProAsp 112
Db 1096 GCGGGTATTGCTTACGATCAAGCGCATCTCGTCATCACGCTAGTCTGCAATTCAGAT 1155
Qy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
Db 1156 ACCGATCGCACCTGGTATAGTTAGGGCAACCTATAAATTCACCCCTAATTTATCTGTT 1215
Qy 133 AspValGlyPheAlaHlsLeuArgGlyLysLysLysHlsPheValGluThrGlnAsnIle 152
Db 1216 GATCTTGGCTATGCTTACTTAAAGGCAAAAGTTTCACTTAAAGAAAGTAAACAAATA 1275
Qy 153 -----LysGlyLeuLeuLeu-----ValGluAlaAspTyrThrThrLysAlaThrAla 168
Db 1276 GGTGCAAAACGTTTCATCTGCATTTAGTACAACTGCAATTTACTTCTCAAGCACACGCA 1335
Qy 169 AsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1336 AATCTTTACGGCTTAAACCTTAAATTTATAGTTTC 1368

```

Search completed: May 19, 2003, 19:39:11
Job time : 1747 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using 'frame_plus.p2n.model'

Run on: May 19, 2003, 18:28:18 ; Search time 172 Seconds
(without alignments)
2343.649 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 OHNGVLGPIYIGKSLTLKP.....ADYTKATANLYGLNLNRYF 179

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlh
-O=/con2.1/USPTO.spool/US0995493/runat_12052003_091308_22995/app_query.fasta_1.327
-DB=N_Geneseq_101002 -GFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -TRANSCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995493 @CGN_1.1.200 @runat_12052003_091308_22995 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	549	57.4	1598	17	AAT28520 H. influenzae dete
2	549	57.4	1598	22	ABA76850 Haemophilus influe
3	549	57.4	1830121	17	AAT42063 Haemophilus influe
4	256	26.8	240	12	AAQ12083 H.influenzae strai
5	151	15.8	1383	21	AAZ54519 Neisseria gonorrhoe
6	144.5	15.1	1329	21	AAA49619 Neisseria meningit
7	144.5	15.1	1401	21	AAZ54520 Neisseria meningit
8	144.5	15.1	65632	21	AAA81502 N. meningitidis pa
C	144.5	15.1	349980	21	AAAF21544 Neisseria meningit
	143.5	15.0	1401	21	AAA49618 Neisseria meningit
	142.5	14.9	1395	21	AAZ54521 Neisseria meningit
	133.5	13.9	1371	24	ABA97441 tmoX nucleotide se
	133.5	13.9	1562	24	ABA97235 tmoX gene nucleoti
	14	11.3	1764	22	AAF25600 H. pylori HP5188 e
	102	10.7	1068	19	AAAX30610 H. pylori outer me
	102	10.7	1764	19	AAAX30611 H. pylori outer me
	98	10.2	1941	22	AAAS43872 Neisseria meningit
	98	10.2	1941	22	AAAD17037 N. meningitidis st
18	96.5	10.1	2127	21	AAZ54328 Neisseria meningit
20	96.5	10.1	56485	21	AAAB1476 N. meningitidis pa
21	96.5	10.1	349980	21	AAAF21612 Neisseria meningit
C	96.5	10.1	837096	21	AAAB1489 N. meningitidis pa
	95.5	10.0	2169	21	AAAS9216 DNA encoding a Nei
24	95	9.9	1746	22	AAAS43869 Neisseria meningit
25	95	9.9	1746	22	AAAD17034 N. meningitidis st
26	94.5	9.9	1719	24	ABQ68105 Listeria monocytog
27	94.5	9.9	1719	24	ABQ69946 Listeria monocytog
28	93.5	9.8	2078	21	AAAS9217 DNA encoding a Nei
29	93	9.7	1035	24	ABAG1422 Haemophilus para
30	93	9.7	1035	24	ABAG1425 Haemophilus para
31	92.5	9.7	465	21	AAZ54327 Neisseria gonorrhe
32	90	9.4	1035	24	ABAG1417 Haemophilus para
33	90	9.4	1035	24	ABAG1418 Haemophilus para
34	90	9.4	1035	24	ABAG1420 Haemophilus para
35	89.5	9.4	1026	24	ABAG1421 Haemophilus para
36	89.5	9.4	1026	24	ABAG1424 Haemophilus para
37	89.5	9.4	1026	24	ABAG1427 Haemophilus para
38	89	9.3	1650	17	AAAT04376 Moraxella catarrha
39	89	9.3	1650	20	AAZ09296 M. catarrhalis out
40	89	9.3	119211	22	AAZ28553 Genomic fragment #
41	88.5	9.2	28730	21	AAAB1486 N. meningitidis pa
42	88	9.2	1035	24	ABAG1419 Haemophilus para
43	88	9.2	1035	24	ABAG1423 Haemophilus para
C	88	9.2	20757	20	AAZ20599 Polynucleotide seq
	87	9.1	1035	24	ABAG1426 Haemophilus para

ALIGNMENTS

RESULT 1
AAT28520

ID AAT28520 standard; DNA; 1598 BP.

XX AAT28520;

AC AAT28520;

XX 02-APR-1997 (first entry)

DT H. Influenzae detection probe #2.

DE H. Influenzae detection probe #2.

XX Detection; probe; amplification primer; bacterial pathogen; pneumonia;

KW Escherichia coli; Klebsiella pneumoniae; pseudomonas aeruginosa;

KW Proteus mirabilis; Streptococcus pneumoniae; Staphylococcus aureus;

KW Staphylococcus epidermidis; Enterococcus faecalis; respiratory tract;

KW Staphylococcus saprophyticus; Streptococcus pyogenes; urinary tract;

KW Haemophilus influenzae; Moraxella catarrhalis; septicaemia; meningitis;

KW infection; intra-abdominal infection; skin infection;

KW bacterial resistance; beta-lactam antibiotic; ds.

CC consists of one or more bacterial colonies. Oligonucleotide
 CC probes and primers complementary to the bacterial genes encoding
 CC resistance to antibiotics such as bla(tem), bla(rob), bla(shv), aadB,
 CC aacC1, aacC2, aacC3, aacC4, mecA, vanA, vanH, vanX, satA, aacA-aphD, vat,
 CC vga, msrA, sul and/or int (ABA76985-ABA77001) are also useful to identify
 CC commonly encountered and clinically important resistance genes. The
 CC invention provides a rapid method of bacterial identification that can be
 CC achieved, which reduces the time currently required for the
 CC identification of pathogens in the clinical laboratory.

XX Sequence 1598 BP; 500 A; 278 C; 329 G; 491 T; 0 other;

Alignment Scores:

Pred. No.: 6,67e-59 Length: 1598
 Score: 549.00 Matches: 105
 Percent Similarity: 76.74% Conservative: 27
 Best Local Similarity: 61.05% Mismatches: 34
 Query Match: 57.37% Indels: 6
 DB: 22 Gaps: 2

US-09-995-493-52 (1-179) x ABA76850 (1-1598)

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 DB 939 AAAGGTAATTAACTTACCTTTACATTGCCAGATTACTTAAAGCTTTCTGGTTTCCATCAATTA 998
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 DB 999 ACTGACAAACTTCGAGTCGATATAGTTATAATATACCATTCGAGTCGTTTAAACAAA 1058
 QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
 DB 1059 TTACATGCCAGCTTCGAAGATGTAATA-----AAAGCTTTTGATAAAGAAATACATAC 1112
 QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrTyrSerLeuAsnAspAlaLeuThrLeu 91
 DB 1113 AGTAATAACTCTCGTTGCTAGGCGCAAGTTATATCTTTATGAAAATTTGACCTTA 1172
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 DB 1173 CGTGGCGGTATTCGTACGATCAAGCGCATCTCGTCATCACCGTAGTGTGCAATTTCCA 1232
 QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 DB 1233 GATACCGATCGCACTTCGGTATAGTTAGGTGCAACCTATATAATTCACGCCGAAATTTATCT 1292
 QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
 DB 1293 GTTGTCTTGGCTATGCTTACTTAAAGCGCAAAAAGTTTCACCTTAAAGAAAGTAAACACA 1352
 QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 DB 1353 ATAGGTGACAAAGGTACATTGACATTGAATACAACTGCAAAATATATCTTCTCAAGCACAC 1412
 QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 DB 1413 GCAAACTTTACGGTTGAATTTAAATTTATAGTTTC 1448

RESULT 3

AAAT42063
 ID AAAT42063 standard; DNA; 1830121 BP.
 XX
 AC AAAT42063;
 XX

DT 14-SEP-1999 (first entry)

XX Haemophilus influenzae complete genome sequence.

XX Genome; bacterium; Haemophilus influenzae; computer readable medium;
 KW expression modulating fragment; regulation; gene expression; vector;
 XX organism; open reading frame; ORF; ds.

OS Haemophilus influenzae.

XX WO9633276-A1.
 PN 24-OCT-1996.
 PD
 XX 22-APR-1996; 96WO-US05320.
 PF
 XX 07-JUN-1995; 95US-0487429.
 PR 21-APR-1995; 95US-0426787.
 PR 07-JUN-1995; 95US-0476102.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
 XX
 DR WPI; 1996-485782/48.
 XX
 PT Haemophilus influenzae Rd genome recorded on computer readable
 PT medium - useful for identifying commercially important nucleic acid
 PT fragments by homology searching

XX Claim 1; Page 77.2-77.1091; 1291pp; English.

CC This sequence represents the complete genome sequence of the bacterium
 CC Haemophilus influenzae strain Rd. The invention relates to a computer
 CC readable medium (CRM) having recorded upon it the complete H-influenzae
 CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
 CC sequence at least 99% identical to (I). By providing the full-length
 CC genomic sequence in a computer readable form, it is possible to identify
 CC commercially important nucleic acid fragments and expression modulating
 CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
 CC regulate the expression of a nucleic acid molecule. Vectors and altered
 CC organisms comprising the predicted ORFs can be used to produce any of the
 CC polypeptide fragments of the H. influenzae Rd genome.

XX Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Alignment Scores:

Pred. No.: 1,26e-54 Length: 1830121
 Score: 549.00 Matches: 106
 Percent Similarity: 77.91% Conservative: 28
 Best Local Similarity: 61.63% Mismatches: 32
 Query Match: 57.37% Indels: 6
 DB: 17 Gaps: 3

US-09-995-493-52 (1-179) x AAAT42063 (1-1830121)

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 DB 422414 AAAGGTGATTAAACCTTACATTTGCCAGATTACTTAGAACCTTTCTGGTTTCCATCAATTA 422473
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 DB 422474 ACTGACAAACTTCGTCGCAATATAGTTATAATATATACCATTCGAGTCGTTTAAACAAA 422533
 QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluTyr 71
 DB 422534 TTAACGCTAGCTCGAAGATGGTAA-----AAAGCTTTTGATAAAGAAATTAACATAC 422587
 QY 72 LysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 DB 422588 AGTAATAACTCTCGTTGCTAGGCGCAAGTTATATCTTTGATGAAAATTTGACCTTA 422647
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 DB 422648 CGTGGCGGTATTCGTACGATCAGCTGCTCGTCATCAGCGTAGTGTGTCANTTTCCA 422707
 QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
 DB 422708 GATACCGATCGCACTTCGGTATAGTTAGGTGCAACCTATATAATTCACGCCGAAATTTATCT 422767
 QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151

Db 1135 TCCGCGGTATGAATACCATATCGGTAAACACACCGTTCGTGATGCCCGCTACACCCAC 1194
 Qy 139 LeuArgGlyLysLysLysHisPheValGluThrGln----- 150
 Db 1195 ATC-----CACAATCAACGACACCGTACCGCGGCGAAGCAAGCGGC 1239
 Qy 151 -----AsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 Db 1240 AACGATGTGGACAGCAAGGC-----GGCTCTTCGCGACGTTTCAAAAACACAC 1287
 Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1288 GCCGACATCATCGGTCTGCAATACACCTACAATTC 1323
 RESULT 7
 ID AA254520 standard; DNA; 1401 BP.
 XX
 AC AA254520;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 989 partial DNA sequence SEQ ID NO:2987.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
 PI Petersen J, Piazza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR P-PSDB; AAY7578.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 7; Page 1404; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254733 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX
 SQ Sequence 1401 BP; 371 A; 457 C; 337 G; 236 T; 0 other;
 Alignment Scores:
 Pred. No.: 4, 15e-08 Length: 1401
 Score: 144.50 Matches: 45
 Percent Similarity: 43.23% Conservative: 38
 Best Local Similarity: 23.44% Mismatches: 70
 Query Match: 15.10% Indels: 39
 DB: 21 Gaps: 7
 US-09-995-493-52 (1-179) x AA254520 (1-1401)
 Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 865 AAAGCCCGGTTAAATCGTTACGCTGTAGTCTTTGTCGTACACGGTATGTACAAGTG 924
 Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe----- 49
 Db 925 TCCGATAAAGCGACCTGTTCGGCGACGTAACCTTGGACGCCACACCGCTTCGATAAG 984
 Qy 50 -----LysGlu-----LeuArgGlyLysTyrGlnAspGlySer 60
 Db 985 GCGGAACGTTGTTTGAAGAAAGAAAAACCGCTCGTCAAGGCAAA----- 1029
 Qy 61 GlyTyrGluAlaPheThrLysLysGluLysLysAspAsnSerArgPheAlaIleGly 80
 Db 1030 ---TCCGACCGCACCCATCACCCCAACTGGCGCAACACCTACAAAGTCGGCTTCGGC 1086
 Qy 81 ThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
 Db 1087 GGTCTTATCAATCAGCGGACCGCTGCAACTGCGCGCGGCGATCGCTTTTGACAATCG 1146
 Qy 101 AlaSerLysThr-----HisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
 Db 1147 CCGTCCCGACCGCGACTACCGCATGACACCTACCGACGCGCAACCGCATCTGGTTC 1206
 Qy 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
 Db 1207 TCCGCGGTATGAATATACCATATCGTAAACACACCGCTCGTATCGCGCTACACCCAC 1266
 Qy 139 LeuArgGlyLysLysLysHisPheValGluThrGln----- 150
 Db 1267 ATC-----CACATCAACGACACCACTCCGCGCGGCAAGCGGC 1311
 Qy 151 -----AsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 Db 1312 AACGATGTGGACAGCAAAAGGC-----GGCTCTTCGCGACGTTTCAAAAACACAC 1359
 Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1360 GCCGACATCATCGGTCTGCAATACACCTACAATTC 1395
 RESULT 8
 ID AA81502 standard; DNA; 65632 BP.
 XX
 AC AA81502;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE N. meningitidis partial DNA sequence gnm_49 SEQ ID NO:49.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200022430-A2.
 XX
 PD 20-APR-2000.
 XX

PF 08-OCT-1999; 99WO-US235573.
 XX
 PR 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 PA (CHIR) CHIRON CORP.
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX WPI; 2000-318079/27.
 DR
 XX
 XX Isolated nucleotide sequences of *Neisseria meningitidis* which can be
 PT used in the diagnosis and treatment of *N. meningitidis* infection and
 PT other *Neisseria* infections, for example, *N. gonorrhoea*.
 XX
 PS Claim 7: Page 1331-1350; 1760pp; English.
 CC The present invention describes methods of obtaining immunogenic
 CC proteins from *Neisseria* genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed *Neisseria meningitidis* genomic DNA
 CC sequences: AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81322 to
 CC AAA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC *Neisseria* bacteria. For example, some of the identified proteins could
 CC be components of vaccines against *Meningococcus B*; against all serotypes;
 CC and/or against all pathogenic *Neisseriae*. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC *Meningococcus B* vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 XX
 SQ Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17126 T; 1 other;

Alignment Scores:
 Pred. No.: 8,98e-06 Length: 65632
 Score: 144.50 Matches: 45
 Percent Similarity: 43.23% Conservative: 38
 Best Local Similarity: 23.44% Mismatches: 70
 Query Match: 15.10% Indels: 39
 DB: 21 Gaps: 7

US-09-995-493-52 (1-179) x AAA81502 (1-65632)

QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuScrGlyPheHisGlnLeu 31
 DB 33852 AAAGCCCGCGTTAAATCGTTCAGCCCTGACTCTTTCGCGTACACGGTATGTACAAAGTC 33911
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe----- 49
 DB 33912 TCCGATAAAGCCGACCTGTTCGCGCAGCTAACTTGCAGCGCCACACGCGCTTCGATAAG 33971
 QY 50 -----LysGlu-----LeuArgGlyLysTyrGlnAspGlySer 60
 DB 33972 GCGGAACCTGTTTTTGAAGAAAGAAAAAACCGTCGTCAAGAGCAAA----- 34016
 QY 61 GlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGly 80
 DB 34017 ---TCCGACCGGCACCACTACCCCACTGCGGCAACACCTACAAAGTCGCTTCGGC 34073

QY 81 ThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
 DB 34074 GGTTCCTTATCAATCAGCAACCGCTGCAACATCGCGCGCGCATCGCTTTTGACAAATCG 34133
 QY 101 AlaSerLysThr-----HisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
 DB 34134 CCGCTCCGCAACGCCGACTACCGCATGAACAGCCTACCGCAGCAACGCGCATCTGTTTC 34193
 QY 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
 DB 34194 TCCGCGCGTATGAAATACCATATCGGTAAACACACGCTGTCGATGCCGCTACACCCAC 34253
 QY 139 LeuArgGlyLysLysLysHisPheValGluThrGln----- 150
 DB 34254 ATC-----CACATCAACGACACCGACGCTACCGCAGCGCAAGCAAGCGGC 34298
 QY 151 -----AsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 DB 34299 AACGATGTGGACAGCAAGGC-----GCGTCTTCGCGACGCTTTCAAAAACCCAC 34346
 QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 DB 34347 GCGGACATCATCGTCTGCAATACACCTACAAATTC 34382

RESULT 9
 ID AAF21544/c
 XX AAF21544 standard; DNA; 349980 BP.
 AC AAF21544;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE *Neisseria meningitidis B* nucleotide sequence SEQ ID NO:1.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO2000066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05928.
 XX
 PR 30-APR-1999; 99US-0132068.
 PR 08-OCT-1999; 99WO-US235573.
 PR 28-FEB-2000; 2000GB-0004695.
 XX
 PA (CHIR) CHIRON CORP.
 XX (GENO-) INST GENOMIC RES.
 PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Massignani V;
 PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
 PI Frazer CM, Grandi G;
 XX
 DR WPI; 2000-647603/62.
 XX
 PT *Neisseria meningitidis B* full length genome sequence and open reading
 PT frames are used to detect, treat and prevent *Neisseria* infections -
 PS
 XX Claim 7: Appendix A; 692pp; English.
 CC The present invention describes the full length genome of
 CC *Neisseria meningitidis B* (NMB). The sequences in AAF21544 and AAF21607-
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8
 CC sequences which overlap each other at the beginning and end of each
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC *Neisseria* proteins given in AAB58550 to AAB58593, and AAF21589 to

CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisseria bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisseria bacteria or of antibodies raised to Neisseria
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.

XX SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other;

Alignment Scores:
 Pred. No.: 9.32e-05 Length: 349980
 Score: 144.50 Matches: 45
 Percent Similarity: 43.23% Conservative: 38
 Best Local Similarity: 23.44% Mismatches: 70
 Query Match: 15.10% Indels: 39
 DB: 21 Gaps: 7

US-09-995-493-52 (1-179) x AAF21544 (1-349980)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 100306 AAGCCCGGTTAAATCGTTCAGCCTGAGTCTTTGTCGCGTACACGGTATGTACAAAGTG 100247
 Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe----- 49
 Db 100246 TCCGATAAAGCCGACCTGTCGGCGCAGCTAACTTGGACGCCACCGCGCTTCGATAAG 100187
 Qy 50 -----LysGlu-----LeuArgGlyLysTyrGlnAspGlySer 60
 Db 100186 CGGAACTGGTTTGTAAAGAGAAACCCGTCGTCAAGGCCAA----- 100142
 Qy 61 GlyTyrGluAlaPheThrLysLysGluTyrLysAspAsnSerArgPheAlaIleGly 80
 Db 100141 ---TCCGACCCGACCCACCATCACCCTGCGGCAACACCTACAAAGTCGGCTTCGGC 100085
 Qy 81 ThrThrTyrSerLeuAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
 Db 100084 GGTTCCTATCAATACAGGACCCGTCGCACTGCGCGCGCGCATCGCTTTGCAAAATCG 100025
 Qy 101 AlaSerLysThr-----HisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
 Db 100024 CCCGTCCGCAACCGCGACTACCGCATGACAGCTACCCGACGCGCACCGCATCTGGTTC 99965
 Qy 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
 Db 99964 TCCGCGGTATGAATAACCATATCCGTAACCAACCCACCGTGTGCGATCGCGCTACACCCAC 99905
 Qy 139 LeuArgGlyLysLysLysHisPheValGluThrGln----- 150
 Db 99904 ATC-----CACATCAACGACACAGCTACCCGCGCGGCAAGCGCGC 99860
 Qy 151 -----AenIleLysLeuLeuValGluAlaAspTyrThrLysAlaThr 167
 Db 99859 AACGATGGGACACCAAGGC-----CGTCTCCGCGACGCTTTCAAAACCCAC 99812
 Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 99811 GCCGACATCGGCTCTGCAATACACCTACAAATTC 99776

RESULT 10

AAA49618

ID AAA49618 standard; DNA; 1401 BP.

XX AAA49618;

AC

XX 27-OCT-2000 (first entry)

XX Neisseria meningitidis BASB044 gene #1.
 DE
 XX
 KW Meningitis; microbial disease; upper respiratory tract infection;
 KW bacteraemia; invasive bacterial disease; BASB044 gene; ds.
 XX
 OS Neisseria meningitidis.
 FH Key Location/Qualifiers
 FT CDS 1..1401
 FT /*tag= a
 FT /product= "BASB044 protein"
 FT sig_peptide 1..72
 FT mat_peptide 73..1401
 FT /*tag= b
 FT /*tag= c
 XX
 PN WC200034482-A2.
 XX
 XX 15-JUN-2000.
 XX
 XX 07-DEC-1999; 99WO-IB02014.
 XX
 XX 08-DEC-1998; 98GB-0026979.
 XX 08-DEC-1998; 98GB-0026980.
 XX 17-DEC-1998; 98GB-0028015.
 XX 03-JAN-1999; 99GB-0000090.
 XX
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PI Ruelle J, Verlant VGCL;
 XX
 XX WPI; 2000-423427/36.
 DR P-PSDB; AAB01258.
 DR
 XX
 XX Novel BASB041, 43, 44 and 48 polypeptides of Neisseria meningitidis
 PT useful for diagnostic, prophylactic and therapeutic purposes against
 PT microbial diseases comprise a specific amino acid sequence -
 PS
 PS Claim 59; Page 163; 171pp; English.
 XX
 CC The present sequence is the BASB044 gene from the Neisseria
 CC meningitidis strain ATCC 13090. The gene and its protein can be used in
 CC the prevention and treatment of microbial disease such as bacteraemia,
 CC meningitis and upper respiratory tract infections. They are particularly
 CC useful for treating bacterial diseases. They can also be used for
 CC diagnosing these diseases. The protein sequence shows significant
 CC homology to the long chain fatty acid transport protein FadL of E. coll.
 XX
 SQ Sequence 1401 BP; 372 A; 458 C; 336 G; 235 T; 0 other;

Alignment Scores:
 Pred. No.: 5.54e-08 Length: 1401
 Score: 143.50 Matches: 45
 Percent Similarity: 43.23% Conservative: 38
 Best Local Similarity: 23.44% Mismatches: 70
 Query Match: 14.99% Indels: 39
 DB: 21 Gaps: 7

US-09-995-493-52 (1-179) x AAA49618 (1-1401)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 865 AAGCCCGGTTAAATCGTTCAGCCTGAGTCTTTGTCGCGTACACGGTATGTACAAAGTG 924
 Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe----- 49
 Db 925 TCCGATAAAGCCGACCTTTCGGCGCAGCTAACTTGGACGCCACCGCGCTTCGATAAG 984
 Qy 50 -----LysGlu-----LeuArgGlyLysTyrGlnAspGlySer 60
 Db 985 CGGAACTGGTTTGTAAAGAGAAACCCGTCGTCAAGGCCAA----- 1029

QY 61 GlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGly 80
 Db 1030 ---TCGACCGCAGCACCACTACCCCACTGGCGCAACACCTACAAAGTGGCTTCGGC 1086
 QY 81 ThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
 Db 1087 GGTCTTATCAATACAGCAACCGCTGCAACTGGCGCGGATCGCTTTTACAAATCG 1146
 QY 101 AlaSerLysThr-----HisLeuSerAlaSerIleProAsnThrAspArgMetTrpTyr 118
 Db 1147 CCGGTCGGACCGCACTACCGCATGACACGCTGCCGCGGACCGCATCTGCTTC 1206
 QY 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
 Db 1207 TCCGCGGTATGAAATACCATATCGGTAAAAACACCGCTCGATCGCGCCCTACACCCAC 1266
 QY 139 LeuArgGlyLysLysHisPheValGluThrGln----- 150
 Db 1267 ATC----- 150
 QY 151 -----AsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 Db 1312 AACGATGTGGACGACCAAGGC-----GGCTCTTCGCGACGTTTCAAAACAC 1359
 QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1360 GCCGACATCATCGCGCTGCAATACACCTACAAATTC 1395
 RESULT 11
 AAZ54521
 ID AAZ54521 standard; DNA; 1395 BP.
 XX
 AC AAZ54521;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE *Neisseria meningitidis* ORF 989 partial DNA sequence SEQ ID NO:2989.
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy; ds.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO9957280-A2.
 PD
 XX 11-NOV-1999.
 PF 30-APR-1999; 99WO-US09346.
 XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX.
 WI: 2000-062150/05.
 DR P-PSDB; AAY75759.
 XX
 PT Novel *Neisseria* polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 7; Page 1405-1406; 1453pp; English.

XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 CC PCR primers used in the exemplification of the present inventions. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 1395 BP; 380 A; 468 C; 320 G; 227 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.36e-08 Length: 1395
 Score: 142.50 Matches: 45
 Percent Similarity: 42.71% Conservative: 37
 Best Local Similarity: 23.44% Mismatches: 71
 Query Match: 14.89% Indels: 39
 DB: 21 Gaps: 7
 US-09-995-493-52 (1-179) x AAZ54521 (1-1395)
 QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
 Db 859 AAAGCCCGCTTAAATCGTTACGCGCGAGTCTTTGTCGTACACGATGATCAAAAGTG 918
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe----- 49
 Db 919 TCGGACAAAGCCGACCTGTTCCGCGACGTAACCTGGACGCGGCACACGCTTCGATAAG 978
 QY 50 -----LysGlu-----LeuArgGlyLysTyrGlnAspGlySer 60
 Db 979 GCGGAACCTGTTTGAAGAAAGAAAGAAACCATCGTCAACGGGCAAA----- 1023
 QY 61 GlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGly 80
 Db 1024 ---TCGACCGCAGCACCACTACCCCACTGGCGCAACACCTACAAAGTGGCTTCGGC 1080
 QY 81 ThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
 Db 1081 GGTCTTATCAATACAGCAACCGCTGCAACTGGCGCGGATCGCTTTTACAAATCG 1140
 QY 101 AlaSerLysThr-----HisLeuSerAlaSerIleProAsnThrAspArgMetTrpTyr 118
 Db 1141 CCGGTCGGACCGCACTACCGCATGACACGCTGCCGCGGACCGCATCTGCTTC 1200
 QY 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
 Db 1201 TCCGCGGATGAAATACCATATCGGCAAAACACCGATCGTGTGCTCCGCTACACCCAC 1260
 QY 139 LeuArgGlyLysLysHisPheValGluThrGln----- 150
 Db 1261 ATC-----CACATCAACGACACCATCGCGCGGCAAGCAAGCGGC 1305
 QY 151 -----AsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
 Db 1306 AACGATGTGGACGACCAAGGC-----GGCTCTTCGCGACGTTTCAAAACAC 1353
 QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1354 GCCGACATCATCGCGCTGCAATACACCTACAAATTC 1389
 RESULT 12
 ABA97241
 ID ABA97241 standard; DNA; 1371 BP.
 XX
 AC ABA97241;
 XX

DT 18-JUN-2002 (first entry)
 XX Tmox nucleotide sequence from ORF2.1.
 DE PHBA: para-Hydrobenzoate; liquid crystal polymer; LCP;
 KW toluene monooxygenase; TMO; pcu gene; p-cresol; ds.
 KW Pseudomonas mendocina KR-1.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..1371
 FT /tag= a
 FT /product= "tmoX"
 XX
 XX W0200192539-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 22-MAY-2001; 2001WO-US16574.
 XX
 XX 01-JUN-2000; 2000US-0585174.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Ben-Bassat A, Cattermole M, Gatenby AA, Gibson KJ;
 PI Ramos-Gonzales MI, Ramos JL, Sariastani S;
 XX WPI: 2002-171436/22.
 DR P-PSDB: ABB08386.
 XX
 XX New nucleic acid fragments encoding bacterial toluene monooxygenase
 PT enzyme pathway, useful for isolating genes encoding proteins from the
 PT same or other microbial species, and for producing para-hydroxybenzoate
 PT
 XX
 XX Claim 2; Page 94-95; 97pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC bacterial toluene monooxygenase enzyme pathway. The nucleic acid
 CC fragments of the invention may be used to isolate genes encoding
 CC proteins from the same or other microbial species. Bacterial strains
 CC transformed with the p-cresol utilizing (pcu) genes are useful for
 CC producing para-hydroxybenzoate (PHBA) which can be used for
 CC synthesizing liquid crystal polymers (LCP). The current sequence
 CC represents the Tmox nucleotide sequence from ORF2.1.
 XX
 XX Sequence 1371 BP; 337 A; 314 C; 364 G; 356 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 9.74e-07 Length: 1371
 Score: 133.50 Matches: 40
 Percent Similarity: 44.00% Conservative: 26
 Best Local Similarity: 26.67% Mismatches: 75
 Query Match: 13.95% Indels: 9
 DB: 24 Gaps: 4

US-09-995-493-52 (1-179) x ABA97241 (1-1371)
 QY 13 GlySerLeuThrLeuLys-----LeuProAlaTyrTrpGluLeuSerGlyPheHis 29
 DB 871 GGCGATATCCCGTAAAGACTTCGAGATGCCCGCCAGCTGACGCTTCGGCTTGCAT 930
 QY 30 GlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg--- 48
 DB 931 CAATCAACGAGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
 QY 49 PheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
 DB 991 ATGAAGACATCATCTGGGATTTCAATCGCAGTCAGGTGGGATGATATCGAATTACCA 1050
 QY 69 GluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrTrpSerLeuAsnAspAla 88
 DB 1051 CACAACTATCAGGATATACGGTGGCCCTCCATCGCACCCTTACAGAGTATATGACAAG 1110

QY 89 LeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAla 108
 DB 1111 CTAACCTCTCGTGTGATATAGCTATGCGCAACAGCGCTGGACAGTAGCTGATATTG 1170
 QY 109 Ser---IleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThr 127
 DB 1171 CCAGTAATTCAGCTTATTGGAAGAAACACAGCTTCTCTCGTAGCGATTATAGTTTTCAT 1230
 QY 128 Pro-----AsnLeuSerValAspValGlyPheAlaLeuLeuAsgGlyLysLys 143
 DB 1231 AAAAATCAAACTCAATTTGGCGATTCTTTGGCCCTAAAGAGACGTTGAACACACCA 1290
 QY 144 LysHisPheValGluThrGlnAsnIleLys 153
 DB 1291 TCATACCTAAGCGCACCGCAACGTTGCAAG 1320
 RESULT 13
 ABA97235
 ID ABA97235 standard; DNA; 1562 BP.
 XX
 XX ABA97235;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 XX tmoX gene nucleotide sequence and its 5' UTR.
 DE PHBA: para-Hydrobenzoate; liquid crystal polymer; LCP;
 KW toluene monooxygenase; TMO; pcu gene; p-cresol; ds.
 KW Pseudomonas mendocina KR-1.
 OS
 XX
 XX Key Location/Qualifiers
 FH 30..46
 FT misc_feature /tag= a
 FT /note= "TODT motif"
 FT -10_signal 124..128
 FT /tag= b
 FT -35_signal 101..105
 FT /tag= c
 XX W0200192539-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 22-MAY-2001; 2001WO-US16574.
 XX
 XX 01-JUN-2000; 2000US-0585174.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Ben-Bassat A, Cattermole M, Gatenby AA, Gibson KJ;
 PI Ramos-Gonzales MI, Ramos JL, Sariastani S;
 XX WPI: 2002-171436/22.
 XX
 XX New nucleic acid fragments encoding bacterial toluene monooxygenase
 PT enzyme pathway, useful for isolating genes encoding proteins from the
 PT same or other microbial species, and for producing para-hydroxybenzoate
 PT
 XX
 XX Example 6; Page 89-90; 97pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC bacterial toluene monooxygenase enzyme pathway. The nucleic acid
 CC fragments of the invention may be used to isolate genes encoding
 CC proteins from the same or other microbial species. Bacterial strains
 CC transformed with the p-cresol utilizing (pcu) genes are useful for
 CC producing para-hydroxybenzoate (PHBA) which can be used for
 CC synthesizing liquid crystal polymers (LCP). The current sequence
 CC represents the tmoX gene nucleotide sequence and its 5' UTR.
 XX
 XX Sequence 1562 BP; 401 A; 345 C; 396 G; 420 T; 0 other;
 SQ

Alignment Scores:
Pred. No.: 1.17e-06 Length: 1562
Score: 133.50 Matches: 40
Percent Similarity: 44.00% Conservatives: 26
Best Local Similarity: 26.67% Mismatches: 75
Query Match: 13.95% Indels: 9
DB: 24 Gaps: 4

US-09-995-493-52 (1-179) x ABA97235 (1-1562)

Oy 13 GlySerLeuThrLeuLys-----LeuProAlaTyrTrpGluLeuSerGlyPheHis 29
||| : : : : :
Db 1062 GCGCATATCCGGTAAAGACTTCGAGATGCCGCCAGTCTGACGCTTCGGCTTCTCAT 1121
||| : : : : :
Oy 30 GlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg--- 48
||| : : : : :
Db 1122 CAATTCAACGACGTTGGCTGGTCTGCTGATGTCACGCGTCTACTGGAGCGATGTC 1181
||| : : : : :
Oy 49 PheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
: : : : :
Db 1182 ATGGAAGACATCAGTGTGGATTCAATCGCAGTCAGTGGGATTGATATCAATTACCA 1241
||| : : : : :
Oy 69 GluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAla 88
||| : : : : :
Db 1242 CACAACATATCAGGATATTACGGTGGCTCCATCGCAGCCGCTTACAGAGTTAATGACAAG 1301
||| : : : : :
Oy 89 LeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAla 108
||| : : : : :
Db 1302 CTAACTCTTGGTGGTATAGCTATGCGCAACAGCGCTCGACAGTAGGCTGATATG 1361
||| : : : : :
Oy 109 Ser-----ileProAspThrAspArgMetTrpTyrSerileGlyAlaThrTyrLysPheThr 127
||| : : : : :
Db 1362 CCAGTAATTCAGCTTATTGAAGAAACACAGCTTCTCTCGGTAGCGATTATAGTTTGTAT 1421
||| : : : : :
Oy 128 Pro-----AsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLys 143
||| : : : : :
Db 1422 AAAAAATCAAACTCAATTTGGCGATTTCTTTGGCTTAAAGAGAGCTTGAACACACCA 1481
||| : : : : :
Oy 144 LysHisPheValGluThrGlnAsnIleLys 153
||| : : : : :
Db 1482 TCATACCTAAGCGGACCCGCAACAGCTTGAAG 1511
||| : : : : :
RESULT 14
ID AAF25600
AC AAF25600 standard; DNA; 1764 BP.
XX AAF25600;
XX
DT 05-APR-2001 (first entry)
DE H. pylori HPS188 encoding DNA.
XX
KW Microbial infection; antibacterial; Helicobacter pylori infection;
KW vaccine; screening; ds.
XX
OS Helicobacter pylori.
XX
PN WO200073502-A2.
XX
PD 07-DEC-2000.
XX
PF 31-MAY-2000; 2000WO-EP05024.
XX
PR 31-MAY-1999; 99DE-1024965.
PR 17-JUN-1999; 99DE-1027740.
PR 21-JUL-1999; 99DE-1034029.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (CREA-) CREATOGEN GMBH.
XX
P1 Apfel H, Fuchs TM, Gibbs CP, Hueck CJ, Meyer TF;
XX

DR WP1: 2001-049948/06.
DR P-PSDB; AAB46323.

XX
PT Preparing an agent for diagnosis or control of microbial infection,
PT useful particularly against Helicobacter, based on identification of
PT essential genes in defective mutants -
XX
PS Claim 37: Page 272-274; 366pp; German.

CC This invention describes a novel preparation of an agent (A) for
CC detection, prevention and/or treatment of microbial infection by:
CC (i) identifying essential genes (I) and corresponding polypeptides
CC (II); (ii) identifying compounds that are directed against (II) and
CC inactivate the microbe; (iii) testing these for suitability for use; and
CC (iv) formulating selected (A). Identifying essential genes (I) comprises
CC preparation of gene-deficient microorganisms by conditional antisense
CC inhibition (CAI) and/or subtractive recombination mutagenesis (SRM),
CC then determining viability and/or survival of the deficient organisms.
CC The products of the invention have antibacterial activity. (A) (which may
CC be a nucleic acid (Ia), vector or host cell containing (Ia), derived
CC polypeptide (IIa), or fragments, (IIa)-specific antibodies or their
CC treatment or prevention of infection by Helicobacter pylori. Particularly
CC (Ia) and (IIa) are used in DNA, subunit or live vaccines. The method
CC identifies essential genes, including those that have homologs in other
CC species, so identified (A) should have a broad spectrum of activity. Many
CC gene-deficient cells can be screened quickly, in an automated process,
CC and the identified genes can be used for screening without purification.
XX
SQ Sequence 1764 BP; 485 A; 364 C; 423 G; 492 T; 0 other;

Alignment Scores:
Pred. No.: 0.00223 Length: 1764
Score: 108.00 Matches: 49
Percent Similarity: 27.14% Conservatives: 29
Best Local Similarity: 23.33% Mismatches: 78
Query Match: 11.29% Indels: 54
DB: 22 Gaps: 10

US-09-995-493-52 (1-179) x AAF25600 (1-1764)

Oy 6 LeuGlyProTyrIleGly-----LysGlySerLeuThrLeu-----LysLeu 19
||| : : : : :
Db 1186 CTTGGCCCTTCTTTAGGAGCGCTTTTGACTAAGGAGCGCTTGAATATCAATGTTCACTC 1245
||| : : : : :
Oy 20 ProAlaTyrTrpGluLeuSerGlyPheHisGln---LeuThrAspGlnTrpAlaIleHis 38
||| : : : : :
Db 1246 CCCCACCCCTAAGCCTAGCCTACGCCACCAATTTTAAAGACCATTTAAGATAGAG 1305
||| : : : : :
Oy 39 TyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeu----- 52
: : : : :
Db 1306 GGGGTGTTTGAGCGTACCTTTTGGAGTCAAGGAATAAATTTTACTAACCCCTGATTTT 1365
||| : : : : :
Oy 53 ---ArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu--- 70
||| : : : : :
Db 1366 GCGAAGCTACTTACAAGGGCTTGACGGGACGGTGGCTTCACTAGCTCTGAGAGCGCTT 1425
||| : : : : :
Oy 71 -----TyrLys 72
: : : : :
Db 1426 AAAAAATGCTAGGCTTAGCGAATTTTAAAGCGTCAACATCGGGGCTGGCTGGAGA 1485
||| : : : : :
Oy 73 AspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArg 92
||| : : : : :
Db 1486 GACACCAACACCTTTAGATTAGGGTAACTTAC---ATGGGTAAGAGCTTGGCTTTGATG 1542
||| : : : : :
Oy 93 AlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAsp 112
||| : : : : :
Db 1543 GGTGCTATTGATTATGACCAAGCCCAAGCCCAAGAG---GCCAGTAGTATCCAGAT 1599
||| : : : : :
Oy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
||| : : : : :
Db 1600 TCCAATGGCTATACCGTGGCTTTTGGGACTAATAACATATTTTAGG-----GGCTTT 1650
||| : : : : :

D**b** 1522 ACCGGTTCGGTTCGAGTTCGACCAAGCAAAACGGCAGCGTAAATCGACATCACCATCCCC 1581

Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu--- 130
 Db 1582 GTTGCCACCTGCA-----ACCGGTTGCAACACTTTACCGACCACTGAAA 1629
 Qy 131 SerValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGln 150
 Db 1630 TCAGCCGACATC---TTGGATGCCGCCCAATATCCGGACATCCGCTTTGTTTCCACCAAA 1686
 Qy 151 ---AsnIleLysGlyLeuLeuValGluAlaAsp 161
 Db 1687 TTCAACTTCAACGGCAAAACTGGTTCCGTTGAC 1722
 RESULT 18
 AAD17037
 ID AAD17037 standard; DNA; 1941 BP.
 XX AAD17037;
 AC AAD17037;
 DT 29-NOV-2001 (first entry)
 DE N. meningitidis strain 394/98 ORF46.1 delta G287N2-953 fusion DNA.
 KW Heterologous expression; Neisserial protein; open reading frame; ORF;
 KW ORF46.1 delta G287N2-953 fusion protein; ds..
 XX Neisseria meningitidis.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1935
 FT /*tag= a
 FT /product= "N. meningitidis New Zealand strain 394/98
 FT ORF46.1 delta G287N2-953 fusion protein"
 XX WO200164920-A2.
 XX 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-IB00420.
 XX 28-FEB-2000; 2000GB-0004695.
 XX 13-NOV-2000; 2000GB-0027675.
 XX (CHIR-) CHIRON SPA.
 PA Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM;
 PI Pizza M;
 XX WPI: 2001-557776/62.
 DR P-FSDB; AAE10018.
 PT Heterologous expression for the expression of two or more Neisserial
 PT proteins in fused state
 PS Example 2; Page 11; 52pp; English.
 XX
 CC The present invention relates to a method for simultaneous heterologous
 CC expression of two or more Neisserial proteins which are in a fused
 CC state. The method is useful for simultaneous heterologous expression of
 CC two or more Neisserial proteins. A protein that may be unstable or
 CC poorly expressed on its own is assisted by adding a suitable hybrid
 CC partner and commercial manufacture is simplified-only one expression and
 CC purification need to be employed in order to produce two separately-
 CC useful proteins. The present sequence is a DNA encoding
 CC Neisseria meningitidis (serogroup B, New Zealand strain 394/98) ORF46.1
 CC (open reading frame) delta G287N2-953 fusion protein.
 XX
 SQ Sequence 1941 BP; 557 A; 494 C; 528 G; 362 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0462 Length: 1941
 Score: 98.00 Matches: 51
 Percent Similarity: 43.60% Conservative: 24

Best Local Similarity: 29.65% Mismatches: 69
 Query Match: 10.24% Indels: 28
 DB: 22 Gaps: 11
 US-09-995-493-52 (1-179) x AAD17037 (1-1941)
 Qy 3 AsnGlyValLeuGlyProTyrIle-----GlyLysGlySerLeuThrLeuLysLeu--- 19
 Db 1252 AACGGCTTTAAGGGACTTGGACGAAATGGCGCGGGATGTTCCGGAAAGTTTTAC 1311
 Qy 20 ---ProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHis 38
 Db 1312 GCGCCGCGCGGAGAGAGTGGCGGA-----AAA 1341
 Qy 39 TyrSerTyrLysTyrThrGluTrpSerArg-----PheLysGluLeuArgGly---Lys 55
 Db 1342 TACAGCTATCGCCCAACAGATCGCGAAAGGGCGGATTCGGCGTGTTCGCGGCAAAAA 1401
 Qy 56 TyrGlnAspGlySerGlyTyrGluAlaPheThr---LysLysGluGluTyrLysAspAsn 74
 Db 1402 GAGCAGGATGATCGGAGGAGGAGGAGCCACCTACAAAGTGGAGCAATATCACGCCAAC 1461
 Qy 75 SerArgPheAlaIle-----GlyThrTyrSerLeuAsnAspAlaLeuThrLeu 91
 Db 1462 GCGCGTTCGCCATCGACCATTTCAACACACGACACACGTCGCGGTTTTTACGGCTGTG 1521
 Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 Db 1522 ACCGGTTCGTCGAGTTGACCAACGACGCGGTAAATCGACATCACCATCCCC 1581
 Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu--- 130
 Db 1582 GTTGCCAACTGCAA-----AGCGGTTGCAACACTTTACCGACCACTGAAA 1629
 Qy 131 SerValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGln 150
 Db 1630 TCAGCCGACATC---TTCGATGCCGCCCAATATCCGGACATCCGCTTTGTTTCCACCAA 1686
 Qy 151 ---AsnIleLysGlyLeuLeuValGluAlaAsp 161
 Db 1687 TTCAACTTCAACGGCAAAACTGGTTCCGTTGAC 1722
 RESULT 19
 AAZ54328
 ID AAZ54328 standard; DNA; 2127 BP.
 XX AAZ54328;
 AC AAZ54328;
 XX
 DT 21-MAR-2000 (first entry)
 DE
 DE Neisseria meningitidis ORF 760 partial DNA sequence SEQ ID NO:2605.
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX

PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scarlato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR P-PSDB; AAY75566.
 XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 PT Claim 7; Page 1235; 1453pp; English.
 XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoea polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254573 represent
 CC PCR primers used in the exemplification of the present inventions. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 2127 BP; 541 A; 624 C; 554 G; 408 T; 0 other:

Alignment Scores:
 Pred. No.: Length: 2127
 Score: 0.081
 Matches: 44
 Percent Similarity: 96.50
 Conservative: 25
 Best local Similarity: 24.44%
 Mismatches: 70
 Indels: 41
 Gaps: 9
 DB: 21

US-09-995-493-52 (1-179) x AA254328 (1-2127)

QY 6 LeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
 Db 1690 TTTGGGAAACCGGTGATGGAAGGTGTGAGACC-----GAAATC 1728
 QY 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu 45
 Db 1729 AGCGGC-----GCGATGACCGGAATGCGAATCCATCGAGGTACAGTACCTGCAC 1782
 QY 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspLysSerGlyTyrGluAlaPhe 65
 Db 1783 ---AGCCAAATCAAAACCGCTCCATTCGCGGACGAAGGC-----ATCTTC 1827
 QY 66 ThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeu 85
 Db 1828 CTGCTGATGCCCAACACGCGCAAC-----CTGTGGACGACTTACCAAGTT 1875
 QY 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHis 105
 Db 1876 ACGTCGGGCTGACCATCGCGCGCGGTG-----AACCGATGAGCGGCATTACT 1926
 QY 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
 Db 1927 TCATCTCGAGGATACATACGCGGTGTTGCCACGCTTCGATCGGCGCATACCGC 1986
 QY 126 PheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHis 145
 Db 1987 TTTACGCCCAAACTGAAGTGTCAATCAACGCGGACAAATC-----TTCAACCGCCAT 2040
 QY 146 PheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrLys 165
 Db 2041 TACTAC-----GCCCGCGTTCGCGCGGAG 2064

QY 166 AlaThrAlaAsnLeuTyrGly-----LeuAsnLeuAsnTyrArgPhe 179
 Db 2065 AGCACCTTTAACATTCCTCCGCGGAGCGAGCCTGACGCAACCTCGCTTACAGTTT 2124
 RESULT 20
 AAA81476
 ID AAA81476 standard; DNA; 56485 BP.
 XX AAA81476;
 AC AAA81476;
 XX 04-DEC-2000 (first entry)
 DT N. meningitidis partial DNA sequence gnm_24 SEQ ID NO:24.
 DE N. meningitidis partial DNA sequence gnm_24 SEQ ID NO:24.
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX Neisseria meningitidis.
 OS Neisseria meningitidis.
 PN WO2000022430-A2.
 XX 20-APR-2000.
 PD 08-OCT-1999; 99WO-US23573.
 PF 09-OCT-1998; 98US-0103794.
 XX 30-APR-1999; 99US-0132068.
 PR (CHIR) CHIRON CORP.
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizsa M;
 XX WPI; 2000-318079/27.
 DR Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea
 XX
 PS Claim 7; Page 507-524; 1760pp; English.
 XX The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament for in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 XX
 SQ Sequence 56485 BP; 12504 A; 14247 C; 16158 G; 13573 T; 3 other;
 Alignment Scores:
 Pred. No.: 7.93 Length: 56485

Score: 96.50 Matches: 44
Percent Similarity: 38.33% Conservative: 25
Best Local Similarity: 24.44% Mismatches: 70
Query Match: 10.08% Indels: 41
DB: 21 Gaps: 9

US-09-995-493-52 (1-179) x AAF21612 (1-349980)

Qy 6 LeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
||||| : : : : : ||| : : : : :
Db 16495 TTGGCAACCGGTGATGAGGTGTTGAGACC-----GAAATC 16533

Qy 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysThrGlu 45
||||| : : : : : ||| : : : : :
Db 16534 AGCGGC-----GCGATCACACCGCAAAATGCCATGCGAGTTACAGCTACCTGCAC 16587

Qy 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrClnAlaPhe 65
||||| : : : : : ||| : : : : :
Db 16588 ---AGCCAAATCAAAACCGCTCCAAATTCGCGCGCAGCAGGAGC-----ATCTTC 16632

Qy 66 ThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeu 85
||||| : : : : : ||| : : : : :
Db 16633 CTGCTGATGCCCAACACAGCGCAAC-----CTGTGACGACTTACCAAGTT 16680

Qy 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaIleSerLysThrHis 105
||||| : : : : : ||| : : : : :
Db 16681 ACGTCGCGGTGACCATCGCGCGCGGTG-----AACGCGATGAGCGGCTACT 16731

Qy 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
||||| : : : : : ||| : : : : :
Db 16732 TCATCTGCGAGGATACATGCGAGCGGTATCCAGCTTCGATGCGTGGCGGATACCGC 16791

Qy 126 PheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHis 145
||||| : : : : : ||| : : : : :
Db 16792 TTCACGCCCAAACTGAAGCTGCAATCAACGCCGACACATC-----TTCAACCGCCAT 16845

Qy 146 PheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLys 165
||| : : : : : ||| : : : : :
Db 16846 TACTAC-----GCCCGCGTCGCGAGCGAG 16869

Qy 166 AlaThrAlaAsnLeuTyrGly-----LeuAsnLeuAsnTyrArgPhe 179
||| : : : : : ||| : : : : :
Db 16870 AGCACCTTACATTCGCGGTTCGAGCGGCGAGCCTGACGCGCAACCTGCTTACAGTTT 16929

RESULT 21
AAF21612/c
ID AAF21612 standard; DNA: 349980 BP.
XX
AC AAF21612;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
KW Neisseria meningitidis.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
OS Neisseria meningitidis.
XX
XX W0200066791-Al.
XX
XX 09-NOV-2000.
XX
XX 08-MAR-2000; 2000WO-US05928.
XX
XX 30-APR-1999; 99US-0132068.
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR) CHIRON CORP.
XX
XX (GENO-) INST GENOMIC RES.
XX

PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Massignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX WPI: 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -
XX
XX Claim 7: Appendix A: 692pp; English.
XX
XX The present invention describes the full length genome of
XX Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
XX to AAF21613 represent fragments of the NMB genomic sequence, as the
XX sequence was too long to go in a record on its own it was split into 8
XX sequences which overlap each other at the beginning and end of each
XX sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
XX the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
XX the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
XX Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
XX AAF21606 represent PCR primers which are used in the exemplification of
XX the present invention. The NMB genome and fragments from it have
XX antibacterial activity, and can be used in vaccines and gene therapy.
XX Neisseria nucleic acids, proteins and/or antibodies which binds to the
XX proteins can be used in compositions for treating or preventing infection
XX due to Neisserial bacteria or as a diagnostic reagent for detecting the
XX presence of Neisserial bacteria or of antibodies raised to Neisserial
XX bacteria. Computers, computer memory, computer storage medium or computer
XX databases can be used in a search to identify open reading frames (ORFs)
XX or coding sequences within the NMB genome. The DNA sequences provide
XX further opportunities to find antigenic or immunogenic proteins which are
XX more effective in vaccines than the outer membrane proteins currently
XX used.

Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

Alignment Scores:

Pred. No.: 101 Length: 349980
Score: 96.50 Matches: 44
Percent Similarity: 38.33% Conservative: 25
Best Local Similarity: 24.44% Mismatches: 70
Query Match: 10.08% Indels: 41
DB: 21 Gaps: 9

US-09-995-493-52 (1-179) x AAF21612 (1-349980)

Qy 6 LeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
||||| : : : : : ||| : : : : :
Db 125159 TTGGCAACCGGTGATGAGGTGTTGAGACC-----GAAATC 125121

Qy 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysThrGlu 45
||||| : : : : : ||| : : : : :
Db 125120 AGCGGC-----GCGATCACACCGCAAAATGCCATGCGAGTTACAGCTACCTGCAC 125067

Qy 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrClnAlaPhe 65
||||| : : : : : ||| : : : : :
Db 125066 ---AGCCAAATCAAAACCGCTCCAAATTCGCGCGCAGCAGGCG-----ATCTTC 125022

Qy 66 ThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeu 85
||||| : : : : : ||| : : : : :
Db 125021 CTGCTGATGCCCAACACAGCGCAAC-----CTGTGACGACTTACCAAGTT 124974

Qy 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaIleSerLysThrHis 105
||||| : : : : : ||| : : : : :
Db 124973 ACGTCGCGGTGACCATCGCGCGCGGTG-----AACGCGATGAGCGGCTACT 124923

Qy 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
||||| : : : : : ||| : : : : :
Db 124922 TCATCTGCGAGGATACATGCGAGCGGTATGCGAGCTTCGATGCGGCGCATACCGC 124863

Qy 126 PheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHis 145
||||| : : : : : ||| : : : : :
Db 124862 TTCACGCCCAAACTGAAGCTGCAATCAACGCCGACACATC-----TTCAACCGCCAT 124809

```

QY 146 PheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLys 165
:::
Db 124808 TACTAC-----GCCCGCTCGGAGCAG 124785
:::
QY 166 AlaThrAlaAsnLeuTyrGly-----LeuAsnLeuAsnTyrArgPhe 179
:::|||||:::|||||:::|||||:::
Db 124784 AGCACCTTTAACTTCGCGGTTCGGAGCGCAGCCTGACGGCAACCTCGCTTACAGTTT 124725

RESULT 22
AA81489/C
ID AAA81489 standard; DNA: 837096 BP.
XX
AC AAA81489;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
XX WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 629-865; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.

```

```

XX
SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Alignment Scores: 343 Length: 837096
Pred. No.: 96.50 Matches: 44
Percent Similarity: 38.33% Conservative: 25
Best Local Similarity: 24.44% Mismatches: 70
Query Match: 10.08% Indels: 41
DB: 21 Gaps: 9

US-09-995-493-52 (1-179) x AAA81489 (1-837096)
QY 6 LeuGlyProTyrIleGlyLysSerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
|||||:::|||||:::|||||:::
Db 489930 TTGGCAACCGCTGATGGAAGGTGTTGAGACC-----GAAATC 489892
|||||:::|||||:::|||||:::
QY 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu 45
|||||:::|||||:::|||||:::
Db 489891 AGCGGC-----GCGATGACACCGCAATGCAATCCATGCAAGGTACACGTACCTGCAC 489838
|||||:::|||||:::|||||:::
QY 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPhe 65
|||||:::|||||:::|||||:::
Db 489837 ---AGCCAAATCAAAACCGCCTCCAATTCGCGCAGCAAGGC-----ATCTTC 489793
|||||:::|||||:::|||||:::
QY 66 ThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeu 85
|||||:::|||||:::|||||:::
Db 489792 CTGCTGATGCCCCAACACAGCGCAAC-----CTGTGACGACTTACCAAGTT 489745
|||||:::|||||:::|||||:::
QY 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaIleSerLysThrHis 105
|||||:::|||||:::|||||:::
Db 489744 ACGTCCGGCTGACCATCGCGCGCGCGTG-----AACGGCATGCGCGCATTTACT 489694
|||||:::|||||:::|||||:::
QY 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
|||||:::|||||:::|||||:::
Db 489693 TCATCTGCGAGGATACATGCGAGCGGTTCATGCCAGTTTCGATGCGCGCATACCGC 489634
|||||:::|||||:::|||||:::
QY 126 PheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHis 145
|||||:::|||||:::|||||:::
Db 489633 TTCACGCCCAAACTGACCTGCAATCAACCCGCAACATC-----TTCAACCGCCAT 489580
|||||:::|||||:::|||||:::
QY 146 PheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLys 165
:::
Db 489579 TACTAC-----GCCCGCTCGGAGCAGCAG 489556
|||||:::|||||:::|||||:::
QY 166 AlaThrAlaAsnLeuTyrGly-----LeuAsnLeuAsnTyrArgPhe 179
|||||:::|||||:::|||||:::
Db 489555 AGCACCTTTAACTTCGCGGTTCGGAGCGCAGCCTGACGGCAACCTCGCTTACAGTTT 489496
|||||:::|||||:::|||||:::

RESULT 23
AAA59216
ID AAA59216 standard; DNA: 2169 BP.
XX
AC AAA59216;
XX
DT 07-NOV-2000 (first entry)
XX
DE DNA encoding a Neisseria meningitidis BASB053 polypeptide.
XX
KW BASB053; Neisseria meningitidis infection; vaccine; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1..2169
FT /tag= a
FT /product= "BASB053"
XX
PN WO200042193-A1.
XX
PD 20-JUL-2000.
XX
PF 10-JAN-2000; 2000WO-EP00137.

```

```
XX 15-JAN-1999; 99GB-0000959.
PR 28-JAN-1999; 99GB-0001903.
XX (SMIK ) SMITHLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
PI WPI; 2000-476062/41.
DR P-PSDB; AAB07697.
XX
XX New Neisseria meningitidis polypeptide useful for diagnosis of
PT Neisseria infection and for development of vaccines against such
PT Infection -
XX
XX Claim 11; Page 55; 92pp; English.
XX
XX The present sequence encodes a Neisseria meningitidis BASB053
CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
CC for BASB053 may be identified in a biological sample in order to
CC diagnose a Neisseria meningitidis infection in an animal. The BASB053
CC polypeptides and polynucleotides may be used as vaccines, for
CC generating an immune response in an animal. A composition comprising
CC at least one antibody immunospecific for BASB053 may be used to
CC treat humans infected with Neisseria meningitidis.
XX
SQ Sequence 2169 BP; 540 A; 663 C; 568 G; 398 T; 0 other;
Alignment Scores:
Pred. NO.: 0.111 Length: 2169
Score: 95.50 Matches: 45
Percent Similarity: 36.07% Conservative: 21
Best Local Similarity: 24.59% Mismatches: 70
Query Match: 9.98% Indels: 47
DB: 21 Gaps: 9
US-09-995-493-52 (1-179) x AAS43869 (1-1746)
Oy 6 LeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
Db 1732 TTGGCAACCGGTGATGCGGAGCGTTGAGACC-----GAAATC 1770
Oy 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysThrGlu 45
Db 1771 AGCGGC-----GCGGTTACACCGAATGGCAATCCATCGAGTTACAGTATCTGCAC 1824
Oy 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPhe 65
Db 1825 ---AGCCAAATCAAAACCGCTCCAAATTCACGCGACGACG-----ATCTTC 1869
Oy 66 ThrLysLysGluGlyTyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeu 85
Db 1870 CTGCTGATGTCGCAACACACAGCGCAAC-----CTGTGACGACTTACCAAGTT 1917
Oy 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHis 105
Db 1918 ACGCCGAGCTGACCATCGCGGCGAGTG-----AACGCGATGCGCGCATTAATTACT 1968
Oy 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
Db 1969 TCATCTGCGAGGATGATCGACGCGGTTATGCCAGTTCGATGCGATGCGGCATACGCG 2028
Oy 126 PheThrProAsnLeuSerValAspValGly-----Phe 136
Db 2029 TTCACGCCCAAGCTCAAGCTCAAAATCAACGCCGACACATCTTCAACCGCATTAATTACT 2088
Oy 137 AlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu 156
Db 2089 GCGCGCGTCGCGCGC-----GCGAACACCTTTAAACATTTCCCGGT----- 2127
Oy 157 LeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuThrGlyLeuAsnLeuAsn 176
Db 2128 -----TCGGAGCGCCACTGACGACGCGCAACCTGGGT 2157
```

```
Oy 177 TyrArgPhe 179
Db 2158 TACAGTTT 2166
RESULT 24
AAS43869
ID AAS43869 standard; DNA; 1746 BP.
XX
XX AAS43869;
XX
XX 18-DEC-2001 (first entry)
XX
XX Neisseria meningitidis fusion protein delta-G287-953 DNA.
XX
XX Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
XX Neisseria protein.
XX
XX Neisseria meningitidis.
XX Synthetic.
XX
XX WO200164922-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-IB00452.
XX
XX 28-NOV-2000; 2000GB-0004695.
XX 13-NOV-2000; 2000GB-0027675.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Arico MB, Comanducci M, Galeotti C, Masignani V, Guilianini MM;
XX Pizza M;
XX
XX WPI; 2001-582163/65.
XX
XX P-PSDB; AAU27570.
XX
XX Producing heterologous proteins from Neisseria meningitidis and N.
XX gonorrhoeae.
XX
XX Example 15; Page 36-37; 119pp; English.
XX
XX The invention relates to methods for the heterologous expression of
XX Neisseria proteins from Neisseria meningitidis and Neisseria
XX gonorrhoeae. At least one domain in the protein is deleted, e.g. the
XX leader peptide, and may be replaced by a domain from a different protein
XX to make a fusion protein, in order to enhance heterologous expression of
XX Neisseria proteins. Also, a region of a protein, such as a poly-glycine
XX stretch, can be mutated to enhance expression. The proteins used in the
XX processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
XX AAS43868-AAS43905 represent DNA molecules encoding Neisseria proteins
XX and peptide regions of proteins of the invention.
XX
XX Sequence 1746 BP; 501 A; 449 C; 459 G; 337 T; 0 other;
XX
Alignment Scores:
Pred. NO.: 0.0949 Length: 1746
Score: 95.00 Matches: 50
Percent Similarity: 43.60% Conservative: 25
Best Local Similarity: 29.07% Mismatches: 69
Query Match: 9.93% Indels: 28
DB: 22 Gaps: 11
US-09-995-493-52 (1-179) x AAS43869 (1-1746)
Oy 3 AsnGlyValLeuGlyProTyrIle-----GlyLysGlySerLeuThrLeuLysLeu--- 19
Db 1057 AACGCGTTTAAGGCGACTTGGACGGAATAATGCGCGGAGTGTTCGGAAGGTTTAC 1116
Oy 20 ---ProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHis 38
Db 1117 GCGCGCGCGCGAGGAGTGTTCGCGGA-----AAA 1146
```


PN WO200228891-A2.
 XX 11-APR-2002.
 PD 04-OCT-2001; 2001WO-FR03061.
 PF 04-OCT-2000; 2000FR-0012697.
 PR (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA Kunst F, Glaser P;
 XX WPI; 2002-332479/37.
 DR New genomic sequences from *Listeria* species, useful for detection,
 XX treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators
 PT Claim 16; SEQ ID 918; 180pp; French.
 PS The present invention relates to nucleic acid sequences
 XX (ABQ671188-ABQ71212) from *Listeria* sp. The sequences are useful as probes
 CC and primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of *Listeria* (potential therapeutic agents), also for
 CC treating infections by *Listeria*, and are useful as immunogens in
 CC anti-*Listeria* vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1719 BP; 634 A; 286 C; 343 G; 456 T; 0 other;

Alignment Scores:
 Pred. No.: 0.107 Length: 1719
 Score: 94.50 Matches: 57
 Percent Similarity: 39.56% Conservative: 32
 Best Local Similarity: 25.33% Mismatches: 72
 Query Match: 9.87% Indels: 64
 DB: 24 Gaps: 11

US-09-995-493-52 (1-1719) x ABQ68105 (1-1719)

QY 4 GlyValLeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrp 23
 DB 418 GCAATAAAGGATCATATAATGCGCAA---AGTGTAAACAACTTGCGCAATATAGT 474
 QY 24 GluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyr 43
 DB 475 GATTCCAAAGGTGGTATCAATCAACGCAAACTTTGGCAATACCCATCACACAAGAA 534
 QY 44 ThrGluTrpSerArgPheLysGluLeuArg---GlyLysTyrGluAspGlySerGlyTyr 62
 DB 535 TCTTTAGAACACAAATGCCAAANAACCTTAGAATGGCCCAAGTTGGGACTCTAGTTATTAC 594
 QY 63 GluAlaPheThrLysLys-----GluLutYrLysAsp-----AsnSer 75
 DB 595 AAAGGTGCATGCGCGAGAGATCCAAACATACAAAGATGCCACTGCATGGTTACAGGGA 654
 QY 76 ArgPheAlaIleGlyThrTyrSer-----LeuAsn----- 86
 DB 655 CGTTATGCAACGGACAACACATATGCTTCTAAGCTTAATATACGCTAATTTCTTCATATAAT 714
 QY 87 -----Asp 87
 DB 715 TTGACTCAATATGATACCTCTGTACGATACGATTAAACAACAAAAAATGTTTCTCAAGAT 774
 QY 88 AlaLeuThrLeuArgAla-----GlyLeuAlaTyrAspLysAla 100
 DB 775 GCTAAGTAGTTAAAGCAGATGGCGATGGTGTGTTTATTAGTGGAAATTTACATACGTCCTGCA 834

QY 101 AlaserLysThrHisLeuSerAlaserIleProAspThrAspArg----- 1115
 DB 835 GCCAGTGCAGAAAAGTTATCTACTGCGAGCGCTTACACAACTAAAGATGTAAAAATTTTA 894
 QY 116 -----MetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValasp 133
 DB 895 AAAGAAGGCACTACAAGCAGAGACGCTGGTCCAAATTTTCTCAATAATAAGTA--- 951
 QY 134 ValGlyPheAlaHisLeuArgGly-----LysLysLysHisPheValGluThr 149
 DB 952 ATCGGCTGGATGATAAACGCGCATTTGTCTATTATCCAAAAGCAACAATGTAAACAG 1011
 QY 150 GlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsn 169
 DB 1012 CTTAACTTAACAGGTAAATC-----ACTGCTGGATCTACTAAT 1050
 QY 170 -LeuTyrGlyLeu 173
 DB 1051 GGTATTATGGTCTG 1063

RESULT 27
 ABQ69946
 ID ABQ69946 standard; DNA; 1719 BP.
 AC ABQ69946;
 XX 29-AUG-2002 (first entry)
 DE *Listeria* monocytogenes EGDe DNA sequence #158.
 KW Antibacterial; *Listeria*; food contamination; mutational analysis;
 XX infection; ds.
 OS *Listeria* monocytogenes EGDe.
 XX WO200228891-A2.
 PD 11-APR-2002.
 PF 04-OCT-2001; 2001WO-FR03061.
 PR 04-OCT-2000; 2000FR-0012697.
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PI Kunst F, Glaser P;
 XX WPI; 2002-332479/37.
 DR New genomic sequences from *Listeria* species, useful for detection,
 XX treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators
 PT Claim 16; SEQ ID 2759; 180pp; French.
 CC The present invention relates to nucleic acid sequences
 CC (ABQ671188-ABQ71212) from *Listeria* sp. The sequences are useful as probes
 CC and primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of *Listeria* (potential therapeutic agents), also for
 CC treating infections by *Listeria*, and are useful as immunogens in
 CC anti-*Listeria* vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1719 BP; 634 A; 286 C; 343 G; 456 T; 0 other;

Alignment Scores:

Qy 137 AlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu 156
 Db 1998 GCCCGCGTGGCGGC-----CGACACCTTTAACTCCCGGT----- 2036
 Qy 157 LeuValGluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeu 176
 Db 2037 ---TCGGAGCGCAGCGTACGGCA-----AACCTGCGT 2066
 Qy 177 TyrArgPhe 179
 Db 2067 TACAGTTT 2075
 RESULT 29
 ABA91422
 ID ABA91422 standard; DNA; 1035 BP.
 AC ABA91422;
 DT 18-APR-2002 (first entry)
 DE Haemophilus paragallinarum strain 0222 haemagglutinin *hagA* gene.
 KW Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;
 KW *hagA*; gene; ds.
 OS Haemophilus paragallinarum.
 FH Key Location/Qualifiers
 FT CDS 1..1035
 FT /*tag= a
 FT sig_peptide 1..63
 FT /*tag= b
 FT mat_peptide 64..1032
 FT /*tag= c
 WO200204485-A1.
 17-JAN-2002.
 06-JUL-2001; 2001WO-AU00822.
 07-JUL-2000; 2000AU-0008652.
 (UYQU) UNIV QUEENSLAND.
 Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 WPI; 2002-154917/20.
 P-PSDB; AAM50721.
 New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 useful as a vaccine for immunising chickens against coryza caused by
 the species .
 Claim 11; Fig 5; 67pp; English.

The present sequence is that of the coding region of the *hagA* gene
 of Haemophilus paragallinarum strain 0222 (serovar B). The gene
 was isolated from chromosomal DNA by PCR amplification. It
 encodes haemagglutinin (see AAM50721). The invention provides
 recombinant haemagglutinin polypeptides (see AAM50716-27) and
 encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 A, B and C) of *H. paragallinarum*, the causative agent of infectious
 coryza of chickens. The polypeptides are useful in vaccines for
 immunisation against infectious coryza, as are the nucleic acids
 when expressed in attenuated bacteria, especially *Salmonella* or
Mycoplasma (claimed). The recombinant polypeptide is preferably
 the mature protein, or a biologically active fragment, variant or
 derivative, that is capable of eliciting an immune response,
 providing protection against one or more strains of *H.*
paragallinarum in chickens. Also claimed are methods of using the

CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.
 SQ Sequence 1035 BP; 318 A; 186 C; 234 G; 297 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.0816 Length: 1035
 Score: 93.00 Matches: 38
 Percent Similarity: 37.50% Conservative: 31
 Best Local Similarity: 20.65% Mismatches: 71
 Query Match: 9.72% Indels: 44
 DB: 24 Gaps: 7

US-09-995-493-52 (1-179) x ABA91422 (1-1035)

Qy 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
 Db 163 ACATTGCGTAATTCGTAACTATTGGGTGTTCCGAGTACCAAAATTAAGTAACTATTC 222
 Qy 36 AlaIleHisTyrSerTyr-----
 Db 223 GCTGTTGAGCTAGGTATGATGACTTTTGGCGCTGCTAAACGCCGTCAAGCGGTGCAAACT 282
 Qy 42 -----LysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
 Db 283 GTTATAAAATACACAAATCAGCGAGCTCAGCTTAAGCTTAAAGCAAGTATATCCAGTGCTT 342
 Qy 60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
 Db 343 GAAGGATTAGATTGTTATGCTCGCGCTTGGAGCAGCGTTGATTCTGTGATTATAACCA 402
 Qy 74 AsnSerArgPheAlaIleGlyThrThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
 Db 403 ACTAAAGCAGCAGCTCCTTAATCAGACGACGACGACATAGCTTAAAGTTTCTCCAGTATTC 462
 Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 Db 463 GCTGCTGCTTTAGAGTATAAC-----TTACCATCATTAACA 498
 Qy 112 Asp-----ThrAspArgMetTrpTyrSer-----Ile 120
 Db 499 GAACCTGCTACCTGCTGTTGTAATATCAATGAGTAAAGTAGGACGTGTGCAAAAAGAT 558
 Qy 121 GlyAlaThrTyrLysPheThrProAsnLeu---SerValaspValGlyPheAlaHisLeu 139
 Db 559 GGTAGCCGCTGATGATTATACACCAAGCATCGGTTCTGTAACGTGCTGTTTATCTTACCGT 618
 Qy 140 ArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGlu 159
 Db 619 TTTGCTCAAGTGCACCATGTTGTAACCTAAGGTTGTTGCAAAAACATTTTGCATTAAAT 678
 Qy 160 AlaAspTyrThr 163
 Db 679 TCAGATGTTACT 690

RESULT 30

ABA91425

ID ABA91425 standard; DNA; 1035 BP.

AC ABA91425;

DT 18-APR-2002 (first entry)

DE Haemophilus paragallinarum strain H-18 haemagglutinin *hagA* gene.

KW Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;

KW *hagA*; gene; ds.

OS Haemophilus paragallinarum.

FH Key Location/Qualifiers

FT CDS 1..1035

FT /*tag= a

FT /product= "haemagglutinin"
 FT 1..63
 FT /*tag= b
 FT mat_peptide 64..1032
 FT /*tag= c
 FT
 XX WO200204485-A1.
 PN
 PD 17-JAN-2002.
 XX
 PD 06-JUL-2001; 2001WO-AU00822.
 PF
 XX 07-JUL-2000; 2000AU-0008652.
 PR
 XX (UYQU) UNIV QUEENSLAND.
 PA
 XX Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 PI WPI; 2002-154917/20.
 XX P-PSDB; AAM50724.
 DR
 XX New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species
 XX
 PS Claim 11; Flg 5; 67pp; English.
 XX
 CC The present sequence is that of the coding region of the hgaA gene
 CC of Haemophilus paragallinarum strain H-18 (serovar C). The gene
 CC was isolated from chromosomal DNA by PCR amplification. It
 CC encodes haemagglutinin (see AAM50724). The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of H. paragallinarum, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially Salmonella or
 CC Mycoplasma (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response,
 CC providing protection against one or more strains of H.
 CC paragallinarum in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.
 XX
 SQ Sequence 1035 BP; 318 A; 185 C; 233 G; 299 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0816 Length: 1035
 Score: 93.00 Matches: 38
 Percent Similarity: 37.50% Conservative: 31
 Best Local Similarity: 20.65% Mismatches: 71
 Query Match: 9.72% Indels: 44
 DB: 24 Gaps: 7
 US-09-995-493-52 (1-179) x ABA91425 (1-1035)
 QY 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
 Db 163 ACATTCGGTAATTCGTGAACCTATGGGGGTTCGGGTGATTACCAAAATTAATGATACTTC 222
 QY 36 AlaileHisTyrSerTyr-----LysGluGluTyrLysAsp 41.
 Db 223 GCTGTTGAGTAGTGTATGATGACTTTGGGCGTGCTAAACCTCGTCAAGCGGTGAAACT 282
 QY 42 -----LysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
 Db 283 GTTATAAAATACAAATACCGAGCTCACTTAAGCTTAAAGCAAGTTATCCAGTGCTT 342
 QY 60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
 Db 343 GAAGGATTAGATGTTTATGTCGCGTGGGACGCGTTGATTGCTTGTGATTATAACCA 402

QY 74 AsnSerArgPheAlaIleGlyThrThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
 Db 403 ACTAAAGAGCAGCTCCTAATCAGACGACGACAAACATAGCTTAAAGTTTCTCCAGTATTC 462
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
 Db 463 GCTGGTGGTTAGAGTATAAC-----TTACCATCATTTACCA 498
 QY 112 Asp-----ThrAspArgMetTyrTyrSer-----Ile 120
 Db 499 GAACTTGCATTACGCTGTTGAATATCAATGGTAAATAAGTAGGACCTGTGGAAAAAGAT 558
 QY 121 GlyAlaThrTyrLysPheThrProAsnLeu---SerValAspValGlyPheAlaHisLeu 139
 Db 559 GGTAGCGGTGTAGATTATACACCAAGCATCGGTTCTGTAACTGCTGTTTATCTTACCGT 618
 QY 140 ArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGlu 159
 Db 619 TTTGGTCAAAAGTGCACCAAGTGTGTAACCTTAAGGTTGTTGCAAAACAACTTGCATTAAAT 678
 QY 160 AlaAspTyrThr 163
 Db 679 TCAGATGTTACT 690
 RESULT 31
 AAZ54327
 ID AAZ54327 standard; DNA; 465 BP.
 XX
 AC AAZ54327;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria gonorrhoea ORF 760 partial DNA sequence SEQ ID NO:2603.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy; ds.
 OS Neisseria gonorrhoea.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 PI Petersen J, Piazza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR P-PSDB; AAY75565.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 7; Page 1234; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoea polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent

CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX SQ Sequence 465 BP; 118 A; 145 C; 119 G; 83 T; 0 other;

Alignment Scores:

Pred. No.: 0.0308 Length: 465
 Score: 92.50 Matches: 47
 Percent Similarity: 36.07% Conservative: 19
 Best Local Similarity: 25.68% Mismatches: 70
 Query Match: 9.67% Indels: 47
 DB: 21 Gaps: 10

US-09-995-493-52 (1-179) x AA254327 (1-465)

Qy 6 LeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeu 25
 Db 28 TTGGGCAAAACCGGTGATGGAGGCGTTGAGACC-----GAAATC 66
 Qy 26 SerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu 45
 Db 67 AGCGGT-----GCGATTACACCGGAATGGCAATCCATGCGAGGTACAGTATCTGCAC 120
 Qy 46 TrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPhe 65
 Db 121 ---AGCCAAATCAAAACCGCGCCCAATCCAGCGCAGCAGCGC-----ATCTTC 165
 Qy 66 ThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeu 85
 Db 166 CTGCTGTGTCACCAACACAGCGGCAAC-----CTGTGACGACTTACCAAGTT 213
 Qy 86 AsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaScrLysThrHis 105
 Db 214 ACGCCCGGCGTACCGTGGCGGGCGGTG-----AACGCGATGCGCGCATTA 264
 Qy 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
 Db 265 TCATCTGCAGGGATGCATGCGAGCGGTTATGCCAGTTTCGATGCGCGCATACCGC 324
 Qy 126 PheThrProAsnLeuSerValAspValGly-----Phe 136
 Db 325 TTCAGCCCAAGCTGAAGCTGCAATCAATGCCGCAACATCTTCACCCGCCATTACTAC 384
 Qy 137 AlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu 156
 Db 385 GCCCGCGTGGCGGCG-----ACGAACACCTTTAATATCCCGGT----- 423
 Qy 157 LeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsn 176
 Db 424 ---TCGGAGCGCAGCGCTACGCGCA-----AACCTGCGGT 453
 Qy 177 TyrArgPhe 179
 Db 454 TACAGTTT 462

RESULT 32

ABA91417
 ID ABA91417 standard; DNA; 1035 BP.

XX AC ABA91417;

XX DT 18-APR-2002 (first entry)

XX DE Haemophilus paragallinarum strain 0083 haemagglutinin in *haga* gene.
 XX

KW Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;
 KW *haga*; gene; ds.
 XX
 OS Haemophilus paragallinarum.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1035
 FT /*tag= a
 FT /product= "haemagglutinin"
 FT sig_peptide 1..63
 FT /*tag= b
 FT mat_peptide 64..1032
 FT /*tag= c
 XX
 PN WO200204485-A1.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WQ-AU00822.
 XX
 PR 07-JUL-2000; 2000AU-0008652.
 XX
 PA (UYQU) UNIV QUEENSLAND.
 XX
 PI Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 XX
 DR WPI; 2002-154917/20.
 DR P-PSDB; AAM50716.
 XX
 PT New haemagglutinin polypeptide isolated from *Haemophilus paragallinarum*
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species .
 XX
 PS Claim 11; Fig 5; 67pp; English.
 XX

CC The present sequence is that of the coding region of the *haga* gene
 CC of *Haemophilus paragallinarum* strain 0083 (serovar A). The gene
 CC was isolated from chromosomal DNA by PCR amplification. It
 CC encodes haemagglutinin (see AAM50716). The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of *H. paragallinarum*, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially *Salmonella* or
 CC *Mycoplasma* (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response.
 CC providing protection against one or more strains of *H.*
 CC *paragallinarum* in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.

XX SQ Sequence 1035 BP; 315 A; 184 C; 235 G; 301 T; 0 other;

Alignment Scores:

Pred. No.: 0.194 Length: 1035
 Score: 90.00 Matches: 37
 Percent Similarity: 36.41% Conservative: 30
 Best Local Similarity: 20.11% Mismatches: 73
 Query Match: 9.40% Indels: 44
 DB: 24 Gaps: 6

US-09-995-493-52 (1-179) x ABA91417 (1-1035)

Qy 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
 Db 163 ACATTGCGTAATCTGTAACTTATGTTGGTGGGTGTTTGGTGTATCAAACTACTGATAATTTT 222
 Qy 36 AlaIleHisTyrSerTyr----- 41
 Db 223 GCTGTTAGCTAGGTTATGATGACTTTGGACGTCGGAACCTTCGTCAGACGGTGCAACT 282

```

Oy 42 -----LysTyrThrGluTyrSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db 283 GTTGGAAACATACAAATCAGGGGCTCATTTAAGCTTAAAGCAAGTATCCAGTCTT 342
Oy 60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
Db 343 GAAGGATTAGATGTTATGCTCGGCTGGAGCTCGGTTAATTCGTTCTGATTATAAACCA 402
Oy 74 AsnSerArgPheAlaIleGlyThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
Db 403 ACTAAAGACAGAGCCCTAATGAGCAGCAGCAACATAGCTTAAAGTTCTCCAGTCTT 462
Oy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 463 GCTGGCTGCCTTAGAGTATAAC-----TTACCATCATTAACCA 498
Oy 112 AspThrAsp-----ArgMetTrpTyrSerIle 120
Db 499 GAATTCGATTACGCTGTTGAATATCAATGGGTGAATAAGTAGGGCGTTGGGAAAAAGAT 558
Oy 121 GlyAlaThrTyrLysPheThrProAsnLeu---SerValAspValGlyPheAlaHisLeu 139
Db 559 GGTAGCGCTGATGATTATACACCAAGCATCGTTCTGTAAGTCTGTTTATCTTACCGT 618
Oy 140 ArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGlu 159
Db 619 TTTGGCCAAAGTGCCACAGTTGTTGAACCTAAGGTTGTTGCAAAACATTTGCATTAAAT 678
Oy 160 AlaAspTyrThr 163
Db 679 TCAGACGTTACT 690

RESULT 33
ID ABA91418
AC ABA91418;
DT 18-APR-2002 (first entry)
DE Haemophilus paragallinarum strain 221 haemagglutinin hgaA gene.
KW Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;
OS Haemophilus paragallinarum.
FH Key Location/Qualifiers
FT CDS 1..1035
FT /tag= a
FT /product= "haemagglutinin"
FT sig_peptide 1..63
FT /tag= b
FT mat_peptide 64..1032
FT /tag= c
XX WO200204485-A1.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-AU00822.
XX 07-JUL-2000; 2000AU-0008652.
XX (UYQU ) UNIV QUEENSLAND.
XX Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX WPI: 2002-154917/20.
XX P-PSDB; AAM50717.
XX New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
XX useful as a vaccine for immunising chickens against coryza caused by

```

```

the species
XX Claim 11; Fig 5; 67pp; English.
XX The present sequence is that of the coding region of the hgaA gene
CC of Haemophilus paragallinarum strain 221 (serovar A). The gene
CC was isolated from chromosomal DNA by PCR amplification. It
CC encodes haemagglutinin (see AAM50717). The invention provides
CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
CC encoding nucleic acids (see ABA91417-27) from 11 strains of infectious
CC A, B and C of H. paragallinarum, the causative agent of coryza in
CC coryza of chickens. The polypeptides are useful in vaccines for
CC immunisation against infectious coryza, as are the nucleic acids
CC when expressed in attenuated bacteria, especially Salmonella or
CC Mycoplasma (claimed). The recombinant polypeptide is preferably
CC the mature protein, or a biologically active fragment, variant or
CC derivative, that is capable of eliciting an immune response,
CC providing protection against one or more strains of H.
CC paragallinarum in chickens. Also claimed are methods of using the
CC haemagglutinin polypeptides and nucleic acids for detection and
CC diagnosis of infectious coryza in chickens.
XX
SQ Sequence 1035 BP; 315 A; 184 C; 235 G; 301 T; 0 other;

Alignment Scores:
Pred. No.: 0.194 Length: 1035
Score: 90.00 Matches: 37
Percent Similarity: 36.41% Conservative: 30
Best Local Similarity: 20.11% Mismatches: 73
Query Match: 9.40% Indels: 44
DB: 24 Gaps: 6

US-09-995-493-52 (1-179) x ABA91418 (1-1035)
Oy 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
Db 163 ACATTGCGTAATCTCTGAACCTATGTTGGGTGTTTGGTATCAATATTACTGATAATTT 222
Oy 36 AlaIleHisTyrSerTyr----- 41
Db 223 GCTGTTGAGTAGGTTTATGATGACTTTGGAGCTGCGAACTTCCTCAAGCGGTGAACCT 282
Oy 42 -----LysTyrThrGluTyrSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db 283 GTTGGAAACATACAAATCAGGGGCTCATTTAAGCTTAAAGCAAGTATCCAGTCTT 342
Oy 60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
Db 343 GAAGGATTAGATGTTATGCTCGGCTGGAGCTCGGTTAATTCGTTCTGATTATAAACCA 402
Oy 74 AsnSerArgPheAlaIleGlyThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
Db 403 ACTAAAGACAGAGCCCTAATGAGCAGCAGCAACATAGCTTAAAGTTCTCCAGTCTT 462
Oy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 463 GCTGGCTGCCTTAGAGTATAAC-----TTACCATCATTAACCA 498
Oy 112 AspThrAsp-----ArgMetTrpTyrSerIle 120
Db 499 GAATTCGATTACGCTGTTGAATATCAATGGGTGAATAAGTAGGGCGTTGGGAAAAAGAT 558
Oy 121 GlyAlaThrTyrLysPheThrProAsnLeu---SerValAspValGlyPheAlaHisLeu 139
Db 559 GGTAGCGCTGATGATTATACACCAAGCATCGTTCTGTAAGTCTGTTTATCTTACCGT 618
Oy 140 ArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGlu 159
Db 619 TTTGGCCAAAGTGCCACAGTTGTTGAACCTAAGGTTGTTGCAAAACATTTGCATTAAAT 678
Oy 160 AlaAspTyrThr 163
Db 679 TCAGACGTTACT 690

```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similar:
Query Match:
DB:

PR 07-JUL-2000; 2000AU-0008652.
 XX (UYQU) UNIV QUEENSLAND.
 PA Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 PI WPI; 2002-154917/20.
 DR P-PSDB; AAM50720.
 XX New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species
 XX Claim 11; Fig 5; 67pp; English.
 PS The present sequence is that of the coding region of the hgaA gene
 CC of Haemophilus paragallinarum strain HD14 (serovar A). The gene
 CC was isolated from chromosomal DNA by PCR amplification. It
 CC encodes haemagglutinin (see AAM50720). The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of H. paragallinarum, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially Salmonella or
 CC Mycoplasma (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response,
 CC providing protection against one or more strains of H.
 CC paragallinarum in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.
 XX Sequence 1026 BP; 316 A; 188 C; 228 G; 294 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 0.222 Length: 1026
 Score: 89.50 Matches: 37
 Percent Similarity: 38.12% Conservative: 32
 Best Local Similarity: 20.44% Mismatches: 71
 Query Match: 9.35% Indels: 41
 DB: 24 Gaps: 7

US-09-995-493-52 (1-179) x ABA91421 (1-1026)

QY 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
 DB 163 ACATTGCGTAATTCGTAACCTATGCTGCTGTTGGTGGTTACCAATTAATCTGATACATTC 222
 QY 36 AlaIleHisTyrSerTyr----- 41
 DB 223 GCTGTGAGCTAGGTATGATGACTTTGGACGTGCGAAATTCGCGCAAGCGGTGAACCT 282
 QY 42 -----LysTyrThrCluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
 DB 283 GTTATAAAACATACAAATCAGCGGGCTCCTTAAGCTTAAAGCTTAAAGCAAGTTATCCAGTGT 342
 QY 60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
 DB 343 GAGGGGTAGATGTTTATGCTCGCTGGGAGCAGCATTCGTTCTGATTATAAACCA 402
 QY 74 AsnSerArgPheAlaIleGlyThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
 DB 403 ACTAAAGAGCAGCCCTAATGAGACGCGACACAGCTTAAAGTATCTCCAGTATT 462
 QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrIleLeuSerAlaSerIlePro 111
 DB 463 GCAGGAGGTTTACAGTATAAC-----TTACCATCATTTACCA 498
 QY 112 Asp-----ThrAspArgMetTrpTyrSer-----IleGlyAlaThr 123
 DB 499 GAACATTGCATTACGTTGTAATCATCAATGGTAAATAAAGTAGGAGCGTGTATGTCGCGT 558

QY 124 TyrLysPheThrProAsnLeu----SerValaspValGlyPheAlaHisLeuArgGlyLys 142
 DB 559 GTAGATTATACACCAAGCATCGGTTCTGTAACTGCTGGTTTACTTACCGTTTTGGTCAA 618
 QY 143 LysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuLeuGluAlaAspTyr 162
 DB 619 AGTGACCAAGTGTGTGAAGCTAAGTGTGTGCAAAACAACTTTCGATTAAATTCAGATGTT 678
 QY 163 Thr 163
 DB 679 ACT 681
 RESULT 36
 ABA91424
 ID ABA91424 standard; DNA; 1026 BP.
 XX ABA91424;
 AC ABA91424;
 XX 18-APR-2002 (first entry)
 DT Haemophilus paragallinarum strain Modesto haemagglutinin hgaA gene.
 DE Haemophilus paragallinarum strain Modesto haemagglutinin hgaA gene.
 XX Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;
 KW hgaA; gene; ds.
 XX Haemophilus paragallinarum.
 OS
 FH Key Location/Qualifiers
 CDS 1..1026
 FT /*tag= a
 FT /product= "haemagglutinin"
 FT sig_peptide 1..63
 FT /*tag= b
 FT mat_peptide 64..1023
 FT /*tag= c
 XX WO200204485-A1.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-AU00822.
 PF 07-JUL-2000; 2000AU-0008652.
 XX (UYQU) UNIV QUEENSLAND.
 PA Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
 P1 WPI; 2002-154917/20.
 XX P-PSDB; AAM50723.
 DR New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
 PT useful as a vaccine for immunising chickens against coryza caused by
 PT the species
 XX Claim 11; Fig 5; 67pp; English.
 PS The present sequence is that of the coding region of the hgaA gene
 CC of Haemophilus paragallinarum strain Modesto (serovar C). The gene
 CC was isolated from chromosomal DNA by PCR amplification. It
 CC encodes haemagglutinin (see AAM50723). The invention provides
 CC recombinant haemagglutinin polypeptides (see AAM50716-27) and
 CC encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
 CC A, B and C) of H. paragallinarum, the causative agent of infectious
 CC coryza of chickens. The polypeptides are useful in vaccines for
 CC immunisation against infectious coryza, as are the nucleic acids
 CC when expressed in attenuated bacteria, especially Salmonella or
 CC Mycoplasma (claimed). The recombinant polypeptide is preferably
 CC the mature protein, or a biologically active fragment, variant or
 CC derivative, that is capable of eliciting an immune response,
 CC providing protection against one or more strains of H.
 CC paragallinarum in chickens. Also claimed are methods of using the
 CC haemagglutinin polypeptides and nucleic acids for detection and
 CC diagnosis of infectious coryza in chickens.

```

CC  diagnosis of infectious coryza in chickens.
XX
SQ  Sequence 1026 BP; 316 A; 188 C; 228 G; 294 T; 0 other;

Alignment Scores:
Pred. No.:      0.222      Length:      1026
Score:          89.50      Matches:      37
Percent Similarity: 38.12%   Conservative: 32
Best Local Similarity: 20.44% Mismatches:  71
Query Match:    9.35%      Indels:      41
DB:             24         Gaps:       7

US-09-995-493-52 (1-179) x ABA91424 (1-1026)
Qy  16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
Db  163 ACATTGCGTAATTCGTAACTATGCTGTTGGTGTACCAAAATACGATAACTTC 222
Qy  36 AlaIleHisTyrSerTyr-----LysGluGluTyrLysAsp 41
Db  223 GCTGTTGAGCTAGCTATATGATGACTTTGGACGTCGAAATTCGCCAAGACGGTGAAACT 282
Qy  42 -----LysTyrThrGluTyrSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db  283 GTTATAAACATACAAATACACGGGCTCATTAAAGCTTAAAGCAAGCTATCCAGTGCTT 342
Qy  60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
Db  343 GAGGGTTAGATGTTTATGCTCGCTTGGACGACGATTCGTTCTCGATTATAACCA 402
Qy  74 AsnSerArgPheAlaIleGlyThrThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91
Db  403 ACTAAAGACGAGCGCCCTTAATGAGACGACGACACACAGCTTAAAGATATCTCCAGTATTT 462
Qy  92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db  463 GCAGGAGGTTAGAGTATAAC-----TTACCATCATTTACCA 498
Qy  112 Asp-----ThrAspArgMetTrpTyrSer-----IleGlyAlaThr 123
Db  499 GAACCTGCATTACGTTGTTGATATCAATGGGCTAAATAAGTAGGACGCTGATGCGGT 558
Qy  124 TyrLysPheThrProAsnLeu---SerValAspValGlyPheAlaHisLeuArgGlyLys 142
Db  559 CTAGATTATACACCAACGATCGGTTCTGTAACCTGCTGCTGTTTATCTATCCGCTTTGGTCAA 618
Qy  143 LysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyr 162
Db  619 AGTCACCCAGTTGTTGAACCTAGGTTGTTGCCAAAACATTTGCCATTAAATTCAGATGTT 678
Qy  163 Thr 163
Db  679 ACT 681

RESULT 37
ABA91427
ID  ABA91427 standard; DNA; 1026 BP.
XX
AC  ABA91427;
XX
DT  18-APR-2002 (first entry)
XX
DE  Haemophilus paragallinarum strain HP60 haemagglutinin hgaA gene.
XX
KW  Haemagglutinin; antigen; vaccine; immunisation; coryza; chicken;
KW  hgaA; gene; ds.
XX
OS  Haemophilus paragallinarum.
XX
FH  key Location/Qualifiers
FT  CDS 1..1026
FT  /tag= a
FT  /product= "haemagglutinin"

```

```

FT  sig_peptide 1..63
FT  mat_peptide /*tag= b
FT  64..1023
FT  /*tag= c
XX
PN  WO200204485-A1.
XX
PD  17-JAN-2002.
XX
PF  06-JUL-2001; 2001WO-AU00822.
XX
PR  07-JUL-2000; 2000AU-0008652.
PA  (UYQU ) UNIV QUEENSLAND.
PI  Terry TD, Tseng H, Hobb RI, Jennings MP, Downes J;
XX
DR  WPI; 2002-154917/20.
XX  P-PSDB; AAM50726.
PT  New haemagglutinin polypeptide isolated from Haemophilus paragallinarum
PT  useful as a vaccine for immunising chickens against coryza caused by
PT  the species
XX
PS  Claim 11; Fig 5; 67pp; English.
XX
CC  The present sequence is that of the coding region of the hgaA gene
CC  of Haemophilus paragallinarum strain HP60 (serovar C). The gene
CC  was isolated from chromosomal DNA by PCR amplification. It
CC  encodes haemagglutinin (see AAM50726). The invention provides
CC  recombinant haemagglutinin polypeptide (see AAM50716-27) and
CC  encoding nucleic acids (see ABA91417-27) from 11 strains (serovars
CC  A, B and C) of H. paragallinarum, the causative agent of infectious
CC  coryza of chickens. The polypeptides are useful in vaccines for
CC  immunisation against infectious coryza, as are the nucleic acids
CC  when expressed in attenuated bacteria, especially Salmonella or
CC  Mycoplasma (claimed). The recombinant polypeptide is preferably
CC  the mature protein, or a biologically active fragment, variant or
CC  derivative, that is capable of eliciting an immune response.
CC  providing protection against one or more strains of H.
CC  paragallinarum in chickens. Also claimed are methods of using the
CC  haemagglutinin polypeptides and nucleic acids for detection and
CC  diagnosis of infectious coryza in chickens.
XX
SQ  Sequence 1026 BP; 316 A; 188 C; 228 G; 294 T; 0 other;

Alignment Scores:
Pred. No.:      0.222      Length:      1026
Score:          89.50      Matches:      37
Percent Similarity: 38.12%   Conservative: 32
Best Local Similarity: 20.44% Mismatches:  71
Query Match:    9.35%      Indels:      41
DB:             24         Gaps:       7

US-09-995-493-52 (1-179) x ABA91427 (1-1026)
Qy  16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrp 35
Db  163 ACATTGCGTAATTCGTAACTATGCTGTTGGTGTACCAAAATACGATAACTTC 222
Qy  36 AlaIleHisTyrSerTyr-----LysGluGluTyrLysAsp 41
Db  223 GCTGTTGAGCTAGCTATATGATGACTTTGGACGTCGAAATTCGCCAAGACGGTGAAACT 282
Qy  42 -----LysTyrThrGluTyrSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db  283 GTTATAAACATACAAATACACGGGCTCATTAAAGCTTAAAGCAAGCTATCCAGTGCTT 342
Qy  60 SerGlyTyrGluAlaPheThrLys-----LysGluGluTyrLysAsp 73
Db  343 GAGGGTTAGATGTTTATGCTCGCTTGGACGACGATTCGTTCTCGATTATAACCA 402
Qy  74 AsnSerArgPheAlaIleGlyThrThr-----TyrSerLeuAsnAspAlaLeuThrLeu 91

```

```

Db 403 ACTAAAGAGAGCCCTAATGAGACGACGAAACAGCTTAAAGATCTCCAGTATT 462
Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 463 GCAGGAGGTTAGATGATTAAC-----TTACCATCATTTACCA 498
Qy 112 Asp-----ThrAspArgMetTrpTyrSer-----IleGlyAlaThr 123
Db 499 GAACTGTCATTACGTTGTAATATCAATGCTAAATAAGTAGGAGCTGATGTCAGCGT 558
Qy 124 TyrLysPheThrProMetLeu---SerValAspValGlyPheAlaHisLeuArgGlyLys 142
Db 559 GTAGATTATATACACCAAGCATCGGTTCTGTAACTGCTGTTATCTTACCGTTTGGTCAA 618
Qy 143 LysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyr 162
Db 619 ACTGCACCAGTTGTTGAACCTAAGTGTGTGCAAAACATTTGCATTAAATTCAGATGTT 678
Qy 163 Thr 163
Db 679 ACT 681

RESULT 38
AAT04376
ID AAT04376 standard; DNA; 1650 BP.
XX
AC AAT04376;
XX
DT 22-JUN-1996 (first entry)
XX
DE Moraxella catarrhalis genomic DNA encoding outer membrane protein E.
XX
KW Bacterium; Branhamella catarrhalis; sinusitis; conjunctivitis;
KW pneumonia; endocarditis; septicemia; meningitis; otitis media;
KW lower respiratory tract infection; chronic bronchitis;
KW chronic obstructive pulmonary disease; vaccine; diagnostic;
KW immunoassay; ds.
XX
OS Moraxella catarrhalis.
XX
FH Key Location/Qualifiers
FT -35_signal 40..45
FT /*tag= a
FT mlsc_feature 51..65
FT /*tag= b
FT /*note= "region of hyphenated dyad symmetry"
FT -10_signal 63..68
FT /*tag= c
FT mlsc_feature 68..90
FT /*tag= d
FT /*note= "region of hyphenated dyad symmetry"
FT RBS 144..148
FT /*tag= e
FT CDS 154..1531
FT /*tag= f
FT /*product= E protein
FT stem_loop 1566..1600
FT /*tag= g
FT /*note= "potential transcriptional terminator"
FT sig_peptide 154..228
FT /*tag= h
XX
PN W09531215-A1.
XX
PD 23-NOV-1995.
XX
PF 20-APR-1995; 95WO-US05134.
XX
PR 17-MAY-1994; 94US-0245758.
XX
PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX

```

```

PI Bhushan R, Murphy TF;
DR WPI: 1996-010692/01.
XX P-PSDB; AAR77432.
PT Vaccine contg outer membrane protein E of Moraxella catarrhalis -
PT for the detection of M.catarrhalis-specific antisera and in
PT diagnostic immunoassays
XX
XX Claim 11; Page 40-43; 58pp; English.
XX
CC This genomic DNA from M. catarrhalis strain 25240 may be used to
CC produce the E protein recombinantly from cells cultured from a host
CC cell system engineered to include a vector containing a nucleotide
CC sequence that regulates expression of DNA encoding E epitopes. The
CC host cell system may be bacteria, yeast, filamentous fungi, insect
CC and mammalian cell lines. The regions of hyphenated dyad symmetry
CC may play a role in regulation of expression of the gene encoding
CC E.
XX
SQ Sequence 1650 BP; 510 A; 316 C; 354 G; 470 T; 0 other;

Alignment Scores:
Pred. No.: 0.498 Length: 1650
Score: 89.00 Matches: 42
Percent Similarity: 33.70% Conservative: 19
Best Local Similarity: 23.20% Mismatches: 84
Query Match: 9.30% Indels: 36
Db: 17 Gaps: 6

US-09-995-493-52 (1-179) x AAT04376 (1-1650)
Qy 9 TyrIleGlyLysCysSerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPhe 28
Db 1027 TATGCGGTAAAGAGATTTTAAAGTTACTTTGCCCTGACATGGAACATTAGATTTTCAA 1086
Qy 29 HisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg 48
Db 1087 ACTGGTGTGAATCAACAACGCTATTAACTGCCAAGTAGTACGTATGTACCATGGTCTGAT 1146
Qy 49 PheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
Db 1147 TTTGACATTCGCCCAACACAGTATACAGAA-----ACCACAAA 1185
Qy 69 GluGluTyr-----LysAspAsnSerArgPheAla 78
Db 1186 CTTGCTTATCCACAGGGTTTACCATCATCATGCTATGACAAAGACCAATGGTCGGTGAA 1245
Qy 79 IleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAsp 98
Db 1246 GTTGGTTTGGTAAGCGTGTAGCGATCGTTGGCTGTTTCAGGTGCGGTAGTTGGGAT 1305
Qy 99 LysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
Db 1306 AGTGGTGCAGGTAAC-----CCTGCAAGTAGCTTAGTCTCATCAAGGCTATTAT 1356
Qy 119 Ser-----IleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPhe 136
Db 1357 TCATGGGGCTTAGTGGCGCGGTATATGTTACACCTGAATGGTGGTCTCTTTGGGT--- 1413
Qy 137 AlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu 156
Db 1414 -----GGTAATACTTTAATTTGGAGATGCTCAA----- 1443
Qy 157 LeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeu 176
Db 1444 -----GCACAGCTACCAACCAAGATAAGTAGTAGTAACCTTTGATAGTAATGATGCG 1494
Qy 177 Tyr 177
Db 1495 TAT 1497
RESULT 39

```


CC The present invention relates to a *Moraxella catarrhalis* genomic library
 CC comprising of a combination of 41 nucleic acid molecules (see
 CC AAF28514-AAF28554). The library has a number of uses described in the
 CC specification e.g. is useful for identifying diagnostic and therapeutic
 CC compositions. *M. catarrhalis* (*Branhamella catarrhalis*) is a large
 CC aerobic, gram-negative diplococcus, normally found among the bacterial
 CC flora of human upper airways. *M. catarrhalis* is known to cause acute,
 CC localised infections such as otitis media, sinusitis and bronchopulmonary
 CC infection and life-threatening, systemic diseases including endocarditis
 CC and meningitis.

XX.
 SQ Sequence 119211 BP; 34729 A; 23773 C; 25664 G; 35041 T; 4 other;

Alignment Scores:

Pred. No.:	198	Length:	119211
Score:	89.00	Matches:	42
Percent Similarity:	33.70%	Conservative:	19
Best Local Similarity:	23.20%	Mismatches:	84
Query Match:	9.30%	Indels:	36
DB:	22	Gaps:	6

US-09-995-493-52 (1-179) x AAF28553 (1-119211)

QY	9	TyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPhe	28
DB	25749	TATGCGGGTAAAGGATTTAAAGTTACTTTGCCCTGACTCATGGAACCTAGATTTTCAA	25808
QY	29	HisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg	48
DB	25809	ACTGGTGTGAATCCCAACACGGCTATTACTGCCAAGTACGCTATGTACCATGGTCTGAT	25868
QY	49	PheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys	68
DB	25869	TTTGACATTCGCGCCCAACACAGTATACAGAA-----ACCACAAA	25907
QY	69	GluGluTyr-----LysAspAsnSerArgPheAla	78
DB	25908	CTTCGTTATCCACAGGGTTTACCAATCATCAGCTATGACAAAGACCAATGGTCGGCTGAA	25967
QY	79	IleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAsp	98
DB	25968	GTTGGTTGGGTAACGGTGTAGCGATCGTTTGGCTGTTTCAGGTCCGGTAGGTGGGAT	26027
QY	99	LysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr	118
DB	26028	ACTGGTGCAGGTAAC-----CCTGCAAGTAGCTTAGGTCTATCAAGAGCTATTAT	26078
QY	119	Ser-----IleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPhe	136
DB	26079	TCATTGGGCTTAGGTCCGGTTATAATGTACACCTGAATGGTCGCTGCTTTGGGT---	26135
QY	137	AlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu	156
DB	26136	-----GGTAATACTTTAAATTTGGAGATGCTCAA-----	26165
QY	157	LeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsn	176
DB	26166	-----GCACAGCTACCAACCAAGATAAAGTAGGTAACTTTTCATAGTAAATGATGGC	26216
QY	177	Tyr	177
DB	26217	TAT	26219

Search completed: May 19, 2003, 19:21:57
 Job time : 902 secs

Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1210 GTTATCTGGCTACTTAAAGCGCAAAAGTTCACTTTAAAGAGTAATAACA 1269
Qy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1270 ATAGGTGACAAACGTACATTTGAATTAACAACTGCAAAATTTATCTCTCAAGCACAC 1329
Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1330 GCAATCTTTACGGTTGCAATTTAATTTATAGTTTC 1365

RESULT 3

US-08-476-102A-5
; Sequence 5, Application US/08476102A
; Patent No. 6355450
; GENERAL INFORMATION:
; APPLICANT: Fleischmann, et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus
; Patent No. 6355450

; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Ave.
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,102A
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

; NAME: Marks, Michelle S.
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P2

TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-309-8439
; TELEX: 301-309-8504

INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

ALIGNMENT SCORES:

; PRED. NO.: 1.8e-61
; SCORE: 549.00
; PERCENT SIMILARITY: 77.91%
; BEST LOCAL SIMILARITY: 61.63%
; QUERY MATCH: 57.37%
; DB: 4

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

US-08-476-102A-5

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

; LENGTH: 1381
; MATCHES: 106
; CONSERVATIVE: 28
; MISMATCHES: 32
; INDLS: 6
; GAPS: 3

Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 992 TTAACCGCTAGCTCGAAGATGTAATA-----AAAGCTTTTGTATGAAGAATTACAAATAC 1045
Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1046 AGTAATAACTCTCGTGTTCATTTAGGGCAAGTATAATCTTTGATGCAAAATTTGACCTTA 1105
Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaLaserLysThrHisLeuSerAlaSerIlePro 111
Db 1106 CGTGGGGGTATTCCTACGATCAGGCTGCATCTCATCAGCGTAGTGCTGCAATTCAC 1165
Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1166 GATACCGATCGCACTTGGTATAGTTTAGGGCAACCTATAAATTCACGCCGAATTTATCT 1225
Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1226 GTTGATCTTGGCTATGCTTAAAGGCAAAAGTTCACTTTAAAGAGTAATAACA 1285
Qy 152 Ile-----LysGlyLeuLeuLeuLeuLeu-----ValGluAlaAspTyrThrThrLysAlaThr 167
Db 1286 ATAGGTGACGAACGTTCACTTACATTTAGTAATGCAAAATTTATCTCTCAAGCACAT 1345
Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1346 GCAATCTTTACGGTTTAAACTTAAATTTATAGTTTC 1381

RESULT 4

US-08-472-172-1

; Sequence 1, Application US/08472172

; Patent No. 5985288

; GENERAL INFORMATION:

; APPLICANT: Munson, Jr., Robert S

; APPLICANT: Grass, Susan

; APPLICANT: Chong, Pele Y Y

; APPLICANT: Fahim, Raafat

; APPLICANT: Sia, Charles D Y

; APPLICANT: McVerry, Patrick

; APPLICANT: Klein, Michel

; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Slim & McBurney

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R6

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,172

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/849,411

; FILING DATE: 07-JUL-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Michael I

; REGISTRATION NUMBER: 24973

; REFERENCE/DOCKET NUMBER: 1038-471

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; TELEX: 063-24567 SIMBAS

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1381

; MATCHES: 106

; CONSERVATIVE: 28

; MISMATCHES: 32

; INDLS: 6

; GAPS: 3

; US-08-472-172-1

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

; LENGTH: 1381

```
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-172-1

Alignment Scores:
Pred. No.: 2,146-61 Length: 1560
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservative: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: 2 Gaps: 2

US-09-995-493-52 (1-179) x US-08-472-172-1 (1-1560)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 922 AAAGGTAATTTAACTTACCTTACATGCCAGATTACTTGAACCTTTCTGGTTCCATCAATTA 981
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 982 ACTGACAACTTCGATGCGCATATAGTTATATAATATACCATGGAGTCGTTTAAACAAA 1041
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 1042 TTACATGCCAGCTTCGAAGATGCTAAA-----AAAGCTTTTCATAAAGAATTAACAATAC 1095
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1096 AGTAATAACTCTCGTGTGCAATTAGGGGCAAGTTATATCTTTATGAAAATTTGACCTTA 1155
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1156 CGTGGCGGTATTCCTTACGATCAAGCGGCATCTCGTCATCACCGTAGTCGTGCAATTC 1215
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1216 GATACCGCATCGCATCTGGTATAGTTAGTGCACACCTATAAAATTCACGCGCAATTTATCT 1275
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1276 GTTGATCTGGCTATGCTTACTTTAAAGGCAAAAAGTTCACATTTAAAGAAGTAAAAACA 1335
QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1336 ATAGGTGACAAACGCTACATTCGACATTGAATTAACAACTGCAAAATTTATCTTCTCAAGCAC 1395
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1396 GCAAACTTTTACGGTTTGAATTTAAATTTATAGTTTC 1431

RESULT 5
US-08-743-637B-26
; Sequence 26, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francols J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OUELLETTE & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Haemophilus Influenzae
; US-08-743-637B-26
;
Alignment Scores:
Pred. No.: 2,216-61 Length: 1598
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservative: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: 2 Gaps: 2

US-09-995-493-52 (1-179) x US-08-743-637B-26 (1-1598)
QY 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 939 AAAGGTAATTTAACTTACCTTACATGCCAGATTACTTGAACCTTTCTGGTTCCATCAATTA 998
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 999 ACTGACAACTTCGATGCGCATATAGTTATATAATATACCATGGAGTCGTTTAAACAAA 1058
QY 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 1059 TTACATGCCAGCTTCGAAGATGCTAAA-----AAAGCTTTTCATAAAGAATTAACAATAC 1112
QY 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1113 AGTAATAACTCTCGTGTGCAATTAGGGGCAAGTTATATCTTTATGAAAATTTGACCTTA 1172
QY 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1173 CGTGGCGGTATTCCTTACGATCAAGCGGCATCTCGTCATCACCGTAGTCGTGCAATTC 1232
QY 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1233 GATACCGCATCGCATCTGGTATAGTTAGTGCACACCTATAAAATTCACGCGCAATTTATCT 1292
QY 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
Db 1293 GTTGATCTGGCTATGCTTACTTTAAAGGCAAAAAGTTCACATTTAAAGAAGTAAAAACA 1352
QY 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1353 ATAGGTGACAAACGCTACATTCGACATTGAATTAACAACTGCAAAATTTATCTTCTCAAGCAC 1412
QY 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
```

```

Db 1413 GCAATCTTTACGGTTTGAATTTAAATATAGTTTC 1448
|||||
RESULT 6
US-08-526-840B-26
; Sequence 26, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: QUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OUALES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,840B
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/304,732
; FILING DATE: 12-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Haemophilus influenzae
US-08-526-840B-26

Alignment Scores:
Pred. No.: 2,21e-61 Length: 1598
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservative: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: 3 Gaps: 2

US-09-995-493-52 (1-179) x US-08-526-840B-26 (1-1598)
Oy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
|||||
Db 939 AAGGTAATTTAACTTTACATGGCAGATTACTTAGAACCTTTCGGTTTCCATCAATTA 998
|||||
Oy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
|||||
Db 999 ACTGACAACTTGCATGTCATTATAGTTATAATATACCCATTGGAGTCGTTTACAAAA 1058
|||||
Oy 52 LeuArgGlyLysGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
|||||

```

```

Db 1059 TTACATGCCAGCTTCGAAGATGGTAAA-----AAAGCTTTTGTAAAGAATTACAATAC 1112
Oy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
|||||
Db 1113 ACTAATAACTCTCGTGTTCATAGGGCAAGTTATAATCTTTATGAAAAATTGCCTTA 1172
|||||
Oy 92 ArgAlaGlyLeuAlaTyrAspLysAlaLaserLysThrHisLysSerAlaSerIlePro 111
|||||
Db 1173 CGTGGGGTATTGCTTACCATCAAGCGGCATCTCGTCATCACCGTAGTCTGCAATTCCA 1232
|||||
Oy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
|||||
Db 1233 GATACGATCGCATTTGGTATAGTTAGTGCAACCTATAAAATTCACGGCGAATTTATCT 1292
|||||
Oy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsn 151
|||||
Db 1293 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAAAGTTCCTTTAAAGAAGTAAAAACA 1352
|||||
Oy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
|||||
Db 1353 ATAGGTGACAAAGCTACATTGACATTGAAATCAACCTGCAAAATATATCTTCTCAAGCACAC 1412
|||||
Oy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
|||||
Db 1413 GCAATCTTTACGGTTTGAATTTAAATATAGTTTC 1448
|||||
RESULT 7
US-08-472-172-11
; Sequence 11, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; TITLE OF INVENTION: of Haemophilus influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-172-11

Alignment Scores:
Pred. No.: 1.08e-24 Length: 240
Score: 256.00 Matches: 50
Percent Similarity: 77.92% Conservatlve: 10
Best Local Similarity: 64.94% Mismatches: 13
Query Match: 26.75% Indels: 4
DB: 2 Gaps: 1

US-09-995-493-52 (1-179) x US-08-472-172-11 (1-240)
QY 107 SerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPhe 126
Db 1 AGTGGTCAATTCAGATACCGATCGACTTGGTATAGTTAGTGCAACTATAAATTC 60
QY 127 ThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHisPhe 146
Db 61 ACGCCGAATTTATCTGTGTATCTGGTATCTTAAAGGCAAAAGTTCACTTT 120
QY 147 ValGluThrGlnAsnIleLysGly-----LeuLeuLeuValGluAlaAspTyr 162
Db 121 AAAGAGTAAACAAATAGGTGACAAACGTACATTGACATTGAATGCAACTGCAAAATTAT 180
QY 163 ThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 181 ACTTCTCAGCACACGCAATCTTAGCGTTTGAATTTAAATTTATAGTTTC 231

RESULT 8
US-08-245-758-11
; Sequence 11, Application US/08245758
; Patent No. 5607846
; GENERAL INFORMATION:
; APPLICANT: Murphy, Timothy F.
; APPLICANT: Bhushan, Reva
; TITLE OF INVENTION: Vaccine for Moraxella catarrhalis
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Met Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5-inch, 1.44 Kb storage
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,758
; FILING DATE: 17/05/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 11520.0051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 11 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1650 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: .yes
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: EMBL-3 clone
; ORIGINAL SOURCE:
; ORGANISM: Moraxella catarrhalis
; STRAIN: 25240
; CELL TYPE: bacterium
; FEATURE:
; LOCATION: E gene region, 154-1531
US-08-245-758-11

Alignment Scores:
Pred. No.: 0.0638 Length: 1650
Score: 89.00 Matches: 42
Percent Similarity: 33.70% Conservatlve: 19
Best Local Similarity: 23.20% Mismatches: 84
Query Match: 9.30% Indels: 36
DB: 1 Gaps: 6

US-09-995-493-52 (1-179) x US-08-245-758-11 (1-1650)
QY 9 TyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPhe 28
Db 1027 TATGCGGGTAAAGAGGATTTTAAAGTTACTTTGCTGACTCATGGAACCTTAGATTTTCAA 1086
QY 29 HisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg 48
Db 1087 ACTGGTGTGAATCCACACACGCTATTAACTGCCAAAGTACGCTATGATGCTCTGAT 1146
QY 49 PhelysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
Db 1147 TTTGACATTCGCCCAACACAGTATACAGAA-----ACCACAAA 1185
QY 69 GluGluTyr-----LysAspAsnSerArgPheAla 78
Db 1186 CTTGCTTATCCACAGGGTTTACCACATCATGCTATGACAAAGCAATGGTCGGCTGAA 1245
QY 79 IleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAsp 98
Db 1246 GTTGGTTGGTAAGCGTGTAGCGATCGTTTTCAGTTCAGTGGCGGTAGGTGGGAT 1305
QY 99 LysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTyrTrp 118
Db 1306 AGTGGTGCAGGTAAC-----CCTGCAAGTAGCTTAGGTCTCTATCAAAAGGCTATTAT 1356
QY 119 Ser-----IleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPhe 136
Db 1357 TCATTGGGCTTAGTGGCGGTATATGTTACACCTGAATGGTGGCTGTCTTTGGGT--- 1413
QY 137 AlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeu 156
Db 1414 -----GGTAAATACTTTAAATTTGGAGATGCTCAA----- 1443
QY 157 LeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsn 176
Db 1444 -----GCACAGCTACCAACCAAGATAAAGTAGGTAACTTTGTAGTAAATGATGTC 1494
QY 177 Tyr 177
Db 1495 TAT 1497

RESULT 9
US-08-810-655A-1
; Sequence 1, Application US/08810655A
; Patent No. 5948412
; GENERAL INFORMATION:
; APPLICANT: Murphy, Timothy F.
; TITLE OF INVENTION: Vaccine for Moraxella catarrhalis
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods &
; ADDRESSEE: Goodyear
; STREET: 1800 One Met Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391

```



```

Db 311 CTAGCTAAGACGGAAATTAATGAGAAAMACTACCTTGGTGGTGAAGACCTGCGGGGTGG 370
Qy 111 ProAspThrAspArgMetTtTpTySerIleGlyAlaThrTyTyrLysPheThrProAsnLeu 130
Db 371 ATAACCATGATTCGACTTGGTACTACCTAGATCCAACTGGTATT----- 418
Qy 131 SerValAspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThrGln 150
Db 419 --ATGCAACACAGTGGCAATATCTA---GGTAATAAGTGGTACTACCTCGCTTCCTCA 472
Qy 151 AsnIleLysGlyLeuLeuValGluAlaAspTyr 162
Db 473 GGAGC-AATGCCCACTGGCTGGTATCAGGAAGGTAC 507

RESULT 13
US-08-183-214-1
; Sequence 1, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; ADP-RIBOSYLARGININE HYDROLASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: . 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.signal
; LOCATION: 1..26
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1115
US-08-183-214-1

Alignment Scores:
Pred. No.: 0.22 Length: 1245
Score: 83.50 Matches: 36
Percent Similarity: 40.74% Conservative: 19

```

Best Local Similarity: 26.67% Mismatches: 55
Query Match: 8.73% Indels: 25
DB: 1 Gaps: 7

US-09-995-493-52 (1-179) x US-08-183-214-1 (1-1245)

OY 11 GlyLysGlySerLeuThrLeuLysLeuPro-----AlaTyrTrpGluLeuSerGly 27
||||| : : : : :
Db 615 GGCAAGGGCTGATGGAGTG---CTGCCAGAGGCCAAAGCGTATGTCACCCAGTCAGGC 671
OY 28 PheHsGlnLeuThrAspGlnTrpAlaIleHsLysTyrSerTyrLysTyrThrGluTrpSer 47
: : : : :
Db 672 TAC-----TTTGTGAAGGAAATCTCCAACTGGTCTTCTCGAAGAAAGTGGAA 725
OY 48 ArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLys 67
: : : : :
Db 726 AGTACCTGGAAGTGGAGG---ATTTGGATGGCAAGTCCGCTCCCGTCTCCGCCAA 782
OY 68 LysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSer----- 84
: : : : :
Db 783 CCCTTTGGTGTGAAGGAAAGGATCAGTTCTATATCGAAGTGAGTACTCAGGCTGGGGT 842
OY 85 -----LeuAsnAspAlaLeuThrLeuArgAlaGly 94
: : : : :
Db 843 GCGAGAGTGGACAGCATGCCCATGATGGCTATGATGCCCTC---CTGGCTGCAGGG 899
OY 95 LeuAlaTyrAspLysAlaAlaSerLysThrHsLeuSerAlaSerIleProAspThrAsp 114
: : : : :
Db 900 GATTCCTGGAAGGAGCTCGCACAGAGCCTTTTCCATCGTGGAGACAGTATCCACG 959
OY 115 Arg-----MetTrpTyrSerIleGlyAlaThrTyrLys 125
: : : : :
Db 960 GCCACCATTCGTGGATGCTGGTGGGAGTTATGCACGGTTTCAAA 1004

RESULT 14

US-09-221-017B-35
; Sequence 35, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER.
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA: PP1546
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA: PP2911
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA: PCT/AU98/01023
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROV, GLADYS H
; REGISTRATION NUMBER: 32,430

; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2903
; US-09-221-017B-35

Alignment Scores:
Pred. No.: 0.733 Length: 2903
Score: 83.50 Matches: 33
Percent Similarity: 41.67% Conservatlve: 22
Best Local Similarity: 25.00% Mismatches: 38
Query Match: 8.73% Indels: 39
DB: 4 Gaps: 8

US-09-995-493-52 (1-179) x US-09-221-017B-35 (1-2903)

OY 10 IleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHs 29
: : : : :
Db 1545 CTCGGAAGGGA---TATGCACGGGGCTGGACTTCTCTGGAAGCATCCGCACTGATC 1601
OY 30 GlnLeuThrAspGlnTrpAlaIleHsLysTyrSerTyrLysTyrThrGluTrpSerArgPhe 49
: : : : :
Db 1602 AGATCTTTCGAGCATGG-----TTCCTCTATTCCTATATACC----- 1637
OY 50 LysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
: : : : :
Db 1638 ---GATGCTCGCAGGCAGTAT-----CTCTTTTCCCGCATCAG 1673
OY 70 GluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSer----- 84
: : : : :
Db 1674 GAACGACCGGAT-----TTCGTAGCCCAAGCACACTCTTCCACCGTGTGAAGTAT 1724
OY 85 -----LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAla 96
: : : : :
Db 1725 TGGTGTGCTCCTATCAGCTCATTTCACAGTGTGCTGCTGGCGTCCCGCATGACC 1784
OY 97 Tyr-----AspLysAlaAlaSerLysThrHsLeuSerAlaSerIleProAspThrAsp 114
: : : : :
Db 1785 TACCAGCATCCCAACATAGCTCCCGGCTTATTGTAACGCATCGTACCGGCCAAT--- 1841
OY 115 ArgMetTrpTyrSerIleGlyAlaThrTyrLysPhe 126
: : : : :
Db 1842 -----TTCAGCATGAGTGCCTGTACATTTAC 1868

RESULT 15

US-08-961-527-146/C
; Sequence 146, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 146:
SEQUENCE CHARACTERISTICS:
LENGTH: 11887 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-146

Alignment Scores:
Pred. No.: 5.45 Length: 11887
Score: 83.50 Matches: 42
Percent Similarity: 39.53% Conservative: 26
Best Local Similarity: 24.42% Mismatches: 77
Query Match: 8.73% Indels: 28
DB: 4 Gaps: 7

US-09-995-493-52 (1-179) x US-08-961-527-146 (1-11887)

Qy 8 ProTyrLeuGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGly 27
Db 11558 CCGTTTCCATCTAAAGTAGTACAAATGGCTTACCACAAATGGTATCAGATTAGAAGT 11499
Qy 28 PheHisGlnLeuThrAspGlnTrpAlaLeuHisTyrSerTyrLys-----TyrThr 44
Db 11498 TTTCACAAAG-----TCAGAGTGG-----TACTACTCGATAAAATGGAGTGTACAA 11451
Qy 45 GluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAla 64
Db 11450 GAGTTTGGTTGGTGAACACATTAGAGATTAAACTAAAGACAGAGTGTGGAGAAAGTAC 11391
Qy 65 PheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaLeuGlyThrTyrSer 84
Db 11390 GGGGAAAACGTGAAGATTCAAGACATAAAGAGAGAGCGTTATTATACGAACATTATAC 11331
Qy 85 LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla----- 100
Db 11330 TTTAATCAAAATCATCTCTTAGAGACAGGTGGCTTTATGATCAGTCTAACTGGTATTAT 11271
Qy 101 AlaSerLysThrHisLeuSer-----AlaSerIle 110
Db 11270 CTAGTAGACGGAATAATTAATGGAGAAACACTACCTGGTGGTGAAGACGTGGGGGTGG 11211
Qy 111 ProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu 130
Db 11210 ATAAACGATGATTCGACTTGGTACTACCTAGATCAACAACTGGTATT----- 11163
Qy 131 SerValAspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGln 150
Db 11162 ---ATGCAACAGGTGGCAATATCTA---GCTAATAAGTGGTACTACTCGTTCCTCA 11109
Qy 151 AsnIleLysGlyLeuLeuLeuValGluAlaAspTyr 162
Db 11108 GGAGC-AATGGCCACTGCTGGTATCAGGAAGGTAC 11074

RESULT 16

US-09-453-702B-194/G
Sequence 194, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Weich, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Plinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 194:
SEQUENCE CHARACTERISTICS:
LENGTH: 9057
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 194:
US-09-453-702B-194

Alignment Scores:
Pred. No.: 4.29 Length: 9057
Score: 83.00 Matches: 42
Percent Similarity: 37.02% Conservative: 25
Best Local Similarity: 23.20% Mismatches: 58
Query Match: 8.67% Indels: 56
DB: 4 Gaps: 10

US-09-995-493-52 (1-179) x US-09-453-702B-194 (1-9057)

Qy 33 AspGlnTrpAlaIleHis-----TyrSerTyrLysTyrThrGluTrpSerArgPheLys 50
Db 1897 GACAAATGGCTATCTCGTGGGGATGACTATCAATCACTCACTGCTGATG----- 1844
Qy 51 GluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 1843 ---TTATTGGCTCATAT-----GCCAGGACATTCGCGCCCGACGATGGCGCAATG 1793
Qy 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSer----- 84
Db 1792 TATACGATTCTAAGCACTTCTCGATTGGTCGCTTCTATACCACTATTGGGTGCCAAN 1733
Qy 84 ----- 84

```
Db 1732 CCGAAGTCTAGCTCCGGAAGAACTACGAAACTCAGGAGTACGGTTTTGGGCTGCCTTTTGAT 1673
QY 85 -----LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
Db 1672 GACCTGATGTTGCCAATGATGCTCTGGAATTTAAAGCCAGCTACTTTGATACCAAGCG 1613
QY 101 -----AlaSerLysThrHlsLeuSerAlaSerIle 110
Db 1612 AAGGATTACATCTCCAGACCGTCGATTTTCGCGCGCGCGAGCACTATGCGTATAACGTC 1553
QY 111 ProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu 130
Db 1552 CCGAAGCC---AAAATCTGG---GGCTGGGATGTGATCAGCAAAATATACCACTGATCTG 1499
QY 131 ---SerValAspValGlyPheAlaHlsLeuArgGlyLysLysLysHlsPheValGluThr 149
Db 1498 TTTAGCTGTGATGGCTATACCGTACCGCGCGGANA-----GACACC 1454
QY 150 GluAsnIleLysGlyLeuLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsn 169
Db 1453 GATACCGCGCAATACATCTCCAGCATTAACCCGGAT---ACTGTTACCACTCTGAAT 1397
QY 170 Leu 170
Db 1396 ATT 1394

RESULT 17
US-08-183-214-9
; Sequence 1, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Marla S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; TITLE OF INVENTION: ADP-RIBOSYLARGININE HYDROLASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1103 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 10..1095
US-08-183-214-9
Alignment Scores:
Pred. No.: 0.335 Length: 1103
Score: 81.50 Matches: 37
Percent Similarity: 40.44% Conservative: 18
Best Local Similarity: 27.21% Mismatches: 54
Query Match: 8.52% Indels: 27
DB: 1 Gaps: 8

US-09-995-493-52 (1-179) x US-08-183-214-9 (1-1103)
QY 11 GlyLysGlySerLeuThrLeuLysLeuPro-----AlaTyrTrpGluLeuSerGly 27
Db 598 GGAAGAGGGCTAATGAGGTG---CTGCCTGAAGCCAAAGTACATCACTCAGTCAGGC 654
QY 28 PheHlsGlnLeuThrAspGlnTrpAlaIleHlsTyrSerTyrLysTyrThrGluTrpSer 47
Db 655 TAC-----TTTGTGAAGGAGAATCTTCAACACTGGTCTCTTCGAGAAAGATGGAA 708
QY 48 ArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLys 67
Db 709 AAGTACTGAGCTTAGAGGA---ATTGTTGGACGGCAACTCGGCTCCCGCTCTTCCGCGAG 765
QY 68 LysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerJ----- 84
Db 766 CCCTTTGGTGTGAAGGAAGGATCAGTTCTACATCAGCTGAGCTACTCGGCTGGGGT 825
QY 85 -----LeuAsnAspAlaLeuThrLeuArgAlaGly 94
Db 826 GGCAGCAGCGGACACGATGCCCTCATGATTGCTACGATGCCCTC---CTGGCCGCGGG 882
QY 95 LeuAlaTyrAspLysAlaAlaSerLysThrHlsLeuSerAlaSerIleProAspThrAsp 114
Db 883 GATTCTCTGGAAGGAGCTCGCACACAGAGCCTTCTTCCACGGTGGACACAGTATCCACG 942
QY 115 Arg-----MetTrpTyrSerIleGlyAlaThrTyrLysPhe 126
Db 943 GCCGCCATGCTGCTGGCTGGTGG-----GGAGTTATGTACGGCTTT 984

RESULT 18
US-08-286-767-1
; Sequence 1, Application US/08286767
; Patent No. 5733760
; GENERAL INFORMATION:
; APPLICANT: Lu, Yichen
; APPLICANT: Miller, Samuel I.
; APPLICANT: Killen, Kevin
; TITLE OF INVENTION: SALMONELLA VECTORS ENCODING TRUNCATED
; TITLE OF INVENTION: pagC FUSION PROTEIN, METHOD OF MAKING, AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RONALD I. EISENSTEIN; DIKE, BRONSTEIN.
; ADDRESSEE: ROBERTS & CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,767
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
```

```

; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 44420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-3400
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2319 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 729..1292
US-08-286-767-1

Alignment Scores:
Pred. No.: 1.12 Length: 2319
Score: 81.00 Matches: 49
Percent Similarity: 37.56% Conservativeness: 28
Best Local Similarity: 23.90% Mismatches: 60
Query Match: 8.46% Indels: 69
DB: 11 Gaps: 11

US-09-995-493-52 (1-179) x US-08-286-767-1 (1-2319)
Qy 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla 36
Db 757 TTATTACTACAGCGTTTGGTT-GTAAATGTTGCACAGCGCGATTAACCCCTTTTCC 815
Qy 37 IleHisTyrSer---TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly--- 54
Db 816 GTGGGTATGCACGGTATGCACAAAGTAAAGTTTCAGGATTTCAAAAATATCCGAGGGTA 875
Qy 55 -----LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 876 AATGTGAATACCGTTATGAGGATGACTCTCCGGTA---AGTTTATTTCCTCGCTAAGT 932
Qy 71 TyrLysAspAsnSerArgPheAla----- 78
Db 933 TACTTATATGGACAGACAGCGCTTCCGGTCTGTTCGACCGTGAAGGTATTCATTACCAT 992
Qy 79 -----IleGlyThrThrTyrSerLeuAsnAsp 87
Db 993 GACAAAGTTTTCAGGTGAGTACGGTCTTTAAATGGTTGGCCAGCCTATCGATTGCTGAC 1052
Qy 88 AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHls 105
Db 1053 AATTTTTCGTTATACCGCGTGGCGGGTGTCCGACCGGTAAGGCGACATTTAAAGAACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATCGCGATCTTTTCTAACAAAATTTCTCAAGGAACAGC 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTGCTGGCGCGGGTCTACAGATGAATCCGCTCGAGAATATCTGCTCGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGGAAGCAACATCTCC--- 1247
Qy 155 LeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAACGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179.
Db 1278 GTTGGATACCGTTTC 1292
RESULT 19
US-08-090-526-1

; Sequence 1, Application US/08090526
; Patent No. 559537
; GENERAL INFORMATION:
; APPLICANT: Miller, Samuel I.
; TITLE OF INVENTION: SALMONELLA VIRULENCE GENES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/090,526
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/629,602
; FILING DATE: 12/18/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/192001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-090-526-1

Alignment Scores:
Pred. No.: 1.12 Length: 2320
Score: 81.00 Matches: 49
Percent Similarity: 37.56% Conservativeness: 28
Best Local Similarity: 23.90% Mismatches: 60
Query Match: 8.46% Indels: 69
DB: 11 Gaps: 11

US-09-995-493-52 (1-179) x US-08-090-526-1 (1-2320)
Qy 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla 36
Db 757 TTATTACTACAGCGTTTGGTT-GTAAATGTTGCACAGCGCGATTAACCCCTTTTCC 815
Qy 37 IleHisTyrSer---TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly--- 54
Db 816 GTGGGTATGCACGGTATGCACAAAGTAAAGTTTCAGGATTTCAAAAATATCCGAGGGTA 875
Qy 55 -----LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 876 AATGTGAATACCGTTATGAGGATGACTCTCCGGTA---AGTTTATTTCCTCGCTAAGT 932
Qy 71 TyrLysAspAsnSerArgPheAla----- 78
Db 933 TACTTATATGGACAGACAGCGCTTCCGGTCTGTTCGACCGTGAAGGTATTCATTACCAT 992
Qy 79 -----IleGlyThrThrTyrSerLeuAsnAsp 87
Db 993 GACAAAGTTTTCAGGTGAGTACGGTCTTTAAATGGTTGGCCAGCCTATCGATTGCTGAC 1052
Qy 88 AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHls 105
Db 1053 AATTTTTCGTTATACCGCGTGGCGGGTGTCCGACCGGTAAGGCGACATTTAAAGAACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATCGCGATCTTTTCTAACAAAATTTCTCAAGGAACAGC 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTGCTGGCGCGGGTCTACAGATGAATCCGCTCGAGAATATCTGCTCGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGGAAGCAACATCTCC--- 1247
Qy 155 LeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAACGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179.
Db 1278 GTTGGATACCGTTTC 1292
RESULT 19
US-08-090-526-1
```

Db 1053 AATTTCCTGTTATACGGCTGGCGGTCTCGGCACGGTAAGGCGACATTAAAGAACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATGGCGATCTTTTCTAACAAAATTTCTCCTCAAGGAAACG 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTCCTGGCGCGGGGTGTACAGATGAATCCGCTGGAGAAATATCGTCGTGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysLysPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGGAAGCAACATCTCC-- 1247
Qy 155 LeuLeuValGluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAACGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179
Db 1278 GTTGATACCGTTTC 1292

RESULT 20

US-08-486-715-1
; Sequence 1, Application US/08486715
; Patent No. 5674724
; GENERAL INFORMATION:
; APPLICANT: MILLER, Samuel I.
; TITLE OF INVENTION: SALMONELLA VIRULENCE GENES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P. C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,715
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/090,536
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/629,602
; FILING DATE: 18-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/192005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-486-715-1
Alignment Scores:
Pred. No.: 1.12
Score: 81.00
Percent Similarity: 37.56%
Best Local Similarity: 23.90%
Length: 2320
Matches: 49
Conservative: 28
Mismatch: 60

Query Match: 8.46% Indels: 69
DB: 1 Gaps: 11
US-09-995-493-52 (1-179) x US-08-486-715-1 (1-2320)
Qy 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGluLeuThrAspGlnTrpAla 36
Db 757 TTATTACTACAAGCGTTTGGTT-GTAAATGTTGCACAGCGCGATACTAAGCGCTTTTCC 815
Qy 37 1leHisTyrSer---TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly--- 54
Db 816 GTGGGGTATCACGGTATGCACAAAGTAAAGTTCAGGATTTCAAAAATATCCGAGGGTA 875
Qy 55 -----LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 876 AATGTGAATACCGTTATGAGGATGACTCTCCGGTA---AGTTTATTCTCCTCGCTAAGT 932
Qy 71 TyrLysAspAsnSerArgPheAla----- 78
Db 933 TACTTATATGAGACACAGCGCTTCGGGTCTGTTGAGCTGAAGGTATTTCATTACCAT 992
Qy 79 -----IleGlyThrThrTyrSerLeuAsnAsp 87
Db 993 GACAAGTTTGAAGTGAAGTACGGTCTTTAATGTTGGCGCAGCTATCGATTGCTGAC 1052
Qy 88 AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHis 105
Db 1053 AATTTTTCGTTATACGCGCTCGCGGGTCTCGGCACGGTAAGGCGACATTTAAGAACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATGGCGATCTTTTCTAACAAAATTTCTCCTCAAGGAAACG 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTCCTGGCGCGGGGTGTACAGATGAATCCGCTGGAGAAATATCGTCGTGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysLysPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGGAAGCAACATCTCC-- 1247
Qy 155 LeuLeuValGluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAACGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179
Db 1278 GTTGATACCGTTTC 1292
RESULT 21
US-08-486-719-1
; Sequence 1, Application US/08486719
; Patent No. 5674736
; GENERAL INFORMATION:
; APPLICANT: MILLER, Samuel I.
; TITLE OF INVENTION: SALMONELLA VIRULENCE GENES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P. C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,719
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/090,536
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/629,602
; FILING DATE: 18-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/192005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-486-719-1
Alignment Scores:
Pred. No.: 1.12
Score: 81.00
Percent Similarity: 37.56%
Best Local Similarity: 23.90%
Length: 2320
Matches: 49
Conservative: 28
Mismatch: 60


```
Qy 71 TyrLysAspAsnSerArgPheAla-----178
Db 933 TACTTATATGAGACAGACAGGCTTCGGGGTCTGTTGAGCCTGAAGGTATTTCATTACCAT 992
Qy 79 -----1leGlyThrThrTyrSerLeuAsnAsp 87
Db 993 GACAAGTTTGAAGTGAAGTACGGTCTTTTAAATGTTGGCCAGCCTATCGATTCTCTGAC 1052
Qy 88 AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaLeuSerLysThrHis 105
Db 1053 AATTTTCGTATACGCGCGGTGTCGGCAGGTAAAGGCGACATTTAAAGACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATGCGGATTTCTTCTAACAAAAATTTCTCAAGGAAACG 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTCGCGCGCGGGGTACAGATGAATCCGCTGGAGAATAATCGCTCGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGCAAGCAACATCTCC---1247
Qy 155 LeuLeuLeuValGluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAAACGGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179
Db 1278 GTTCGATACCGTTTC 1292

RESULT 24
US-08-476-100-1
: Sequence 1, Application US/08476100
: Patent No. 5731196
: GENERAL INFORMATION:
: APPLICANT: Miller, Samuel I., III
: APPLICANT: Mekalanos, John J.
: TITLE OF INVENTION: SALMONELLA VIRULENCE GENES
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P. C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: COMPUTER: IBM PS/2 Model 502 or 55SX
: OPERATING SYSTEM: MS-DOS (Version 5.0)
: SOFTWARE: WordPerfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/476,100
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/090,526
: FILING DATE: 09-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/629,602
: FILING DATE: 18-DEC-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Clark, Paul T.
: REGISTRATION NUMBER: 30,162
: REFERENCE/DOCKET NUMBER: 00786/192003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
```

```
: LENGTH: 2320
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-476-100-1

Alignment Scores:
Pred. No.: 1,12 Length: 2320
Score: 81.00 Matches: 49
Percent Similarity: 37.56% Conservative: 28
Best Local Similarity: 23.90% Mismatches: 60
Query Match: 8.46% Indels: 69
DB: 1 Caps: 11

US-09-995-493-52 (1-179) x US-08-476-100-1 (1-2320)
Qy 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla 36
Db 757 TTATTACTACAAGCGTTTGGTT-GTAAATGTTGCACAGCGCGATACTAAGCCCTTTTCC 815
Qy 37 IleHisTyrSer---TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly--- 54
Db 816 GTGGGGTATGCACGGTATGCACAAAGTAAAGTTCAGGATTTCAAAAATATCCGAGGGTA 875
Qy 55 -----LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 876 AATCTGAATACCGTTATGAGGATGACTCTCCGGTA---AGTTTATTTCCTCGCTAAGT 932
Qy 71 TyrLysAspAsnSerArgPheAla-----78
Db 933 TACTTATATGAGACAGACAGGCTTCGGGGTCTGTTGAGCCTGAAGGTATTTCATTACCAT 992
Qy 79 -----IleGlyThrThrTyrSerLeuAsnAsp 87
Db 993 GACAAGTTTGAAGTGAAGTACGGTCTTTTAAATGTTGGCCAGCCTATCGATTCTCTGAC 1052
Qy 88 AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaLeuSerLysThrHis 105
Db 1053 AATTTTCGTATACGCGCGGTGTCGGCAGGTAAAGGCGACATTTAAAGACAT 1112
Qy 106 LeuSerAlaSerIleProAspThrAsp-----ArgMet 116
Db 1113 -----TCCACTCAGGATGCGGATTTCTTCTAACAAAAATTTCTCAAGGAAACG 1163
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal 134
Db 1164 GGATTTCGCGCGCGGGGTACAGATGAATCCGCTGGAGAATAATCGCTCGATGTT 1223
Qy 135 GlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGly 154
Db 1224 GGGTAT-----GAAGCAAGCAACATCTCC---1247
Qy 155 LeuLeuLeuValGluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn 174
Db 1248 -----TCTACAAA-----ATAAACGGGCTTCAACGTCGGG 1277
Qy 175 LeuAsnTyrArgPhe 179
Db 1278 GTTCGATACCGTTTC 1292

RESULT 24
US-08-565-861-1
: Sequence 1, Application US/08565861
: Patent No. 5843426
: GENERAL INFORMATION:
: APPLICANT: Miller, Samuel I.
: APPLICANT: Mekalanos, John J.
: APPLICANT: Hohmann, Elizabeth
: TITLE OF INVENTION: SALMONELLA VACCINES
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
```


DB:	3	Gaps:	11
US-09-995-493-52 (1-179) x US-08-475-749-1 (1-2320)			
Qy	17	LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla	36
Db	757	TTATTACTACAACGGTTTGGTT-GTAAATGTTCACAGGCCGATACTAAGCCCTTTTCC	815
Qy	37	IleHisTyrSer---TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly---	54
Db	816	GTGGGGTATGCACGGTATGCACAAAGTTCAGGATTTCAAAAATATCCGAGGGGTA	875
Qy	55	---LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu	70
Db	876	ANTGTGAAATACCGGTATGAGGATGACTCTCCGGTA---AGTTTATTTTCTCGCTAAGT	932
Qy	71	TyrLysAspAsnSerArgPheAla---	78
Db	933	TACTTATATGCAGACACAGAGCGTTCGGGCTGTGTGAGCCTGAAGGTATTTACCAT	992
Qy	79	-----IleGlyThrThrTyrSerLeuAsnAsp	87
Db	993	GACAAGTTTGAGTGGAAGTACGGTTCTTTAATGGTTGGCGCAGCTATPCGATTCCTCTGAC	1052
Qy	88	AlaLeuThrLeu-----ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHis	105
Db	1053	AAATTTCGTATTACCGCGCTGGCGGGTCTGCCACGGTAAAGCGCACATTTTAAAGACAT	1112
Qy	106	LeuSerAlaSerIleProAspThrAsp-----ArgMet	116
Db	1113	-----TCCACTCAGGATCGCATTTCTTTCTAACAAAATTTCTCTCAAGGAAAACG	1163
Qy	117	TrpTyrSerIleGlyAlaThrTyrLysPheThrPro-----AsnLeuSerValAspVal	134
Db	1164	GGATTTCCTGGCGCGGGGTACAGATGATCCGCTGGAGATATATCGTCTCATGTTT	1223
Qy	135	GlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGly	154
Db	1224	GGGTAT-----GAAGGAAGCAACATCTCC---	1247
Qy	155	LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsn	174
Db	1248	-----TCTACAAA-----ATAAAGCGGCTTCAACGTCGGG	1277
Qy	175	LeuAsnTyrArgPhe	179
Db	1278	GTGGGATACCGTTTC	1292

RESULT 26	
PCT-US94-07658-1	
: Sequence 1, Application PC/TUS9407658	
: GENERAL INFORMATION:	
: APPLICANT: Miller, Samuel I.	
: APPLICANT: Mekalanos, John J.	
: TITLE OF INVENTION: SALMONELLA VACCINES	
: NUMBER OF SEQUENCES: 15	
: CORRESPONDENCE ADDRESS:	
: ADDRESSEE: Fish & Richardson	
: STREET: 225 Franklin Street	
: CITY: Boston	
: STATE: Massachusetts	
: COUNTRY: U.S.A.	
: ZIP: 02110-2804	
: COMPUTER READABLE FORM:	
: MEDIUM TYPE: Floppy disk	
: COMPUTER: IBM PC compatible	
: OPERATING SYSTEM: PC-DOS/MS-DOS	
: SOFTWARE: Patent In Release #1.0, Version	
: SOFTWARE: #1.25 and WordPerfect (Version 5.1)	
: CURRENT APPLICATION DATA:	
: APPLICATION NUMBER: PCT/US94/07658	
: FILING DATE:	
: CLASSIFICATION:	


```

RESULT 27
US-09-370-838-169
; Sequence 169, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 169
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-169

```


QY 52 ---LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 1271 CATATGAAAGCGTCATTCATCCGATGTCAGTGTGCTGTGTGAAGAAAGCCTCC 1330
QY 71 TyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThr 90
Db 1331 TACCTTGATGTCATCAGG---GCCATTGGCGCAAGAACGCG-----GATGCTGTGACA 1381
QY 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaLaSerLysThrHisLeuSer----- 107
Db 1382 CTGGATGTCAGTTTGGTGTATGATCCTTACTTGGCTCCCAATAACCTGAAGCCTGTGGTG 1441
QY 108 -----AlaSerIleProAspThrAspArgMetTyrTyrSerIleGlyAlaThr 123
Db 1442 GCAGAGTTCTATGCGGTCAAAGAGGATCCACAGACTTTCTATTATGCTTGTGTGGTG 1501
QY 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 1502 AAGAAG-----GATAGTGGCTTCCAGATGAACAGCTTCGAGGC 1540
QY 142 LysLys 143
Db 1541 AAGAAG 1546
RESULT 31
US-09-037-188-1
; Sequence 1, Application US/09037188
; Patent No. 6027921
; GENERAL INFORMATION:
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; APPLICANT: Concino, Michael F.
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE
; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,188
; FILING DATE: 02-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Ph.D., J.D., Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 07236/009002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8906
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3427 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA to mRNA
US-09-037-188-1
Alignment Scores: 2.63 Length: 3427
Pred. No.:

Score: 80.00 Matches: 34
Percent Similarity: 40.16% Conservative: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.36% Indels: 20
Db: 3 Gaps: 6
US-09-995-493-52 (1-179) x US-09-037-188-1 (1-3427)
QY 32 ThrAspGlnTyrAlaIleHisTyrSerTyrLysTyrThrGluTyrSerArgPheLysGlu 51
Db 1211 ACTGTGAGATGTCGTCAGTGTCCGAGCATGAGCCACTAAGTGCAGAGTTTCGCGAC 1270
QY 52 ---LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 1271 CATATGAAAGCGTCATTCATCCGATGTCAGTGTGCTGTGTGAAGAAAGCCTCC 1330
QY 71 TyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThr 90
Db 1331 TACCTTGATGTCATCAGG---GCCATTGGCGCAAGAACGCG-----GATGCTGTGACA 1381
QY 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaLaSerLysThrHisLeuSer----- 107
Db 1382 CTGGATGTCAGTTTGGTGTATGATCCTTACTTGGCTCCCAATAACCTGAAGCCTGTGGTG 1441
QY 108 -----AlaSerIleProAspThrAspArgMetTyrTyrSerIleGlyAlaThr 123
Db 1442 GCAGAGTTCTATGCGGTCAAAGAGGATCCACAGACTTTCTATTATGCTTGTGTGGTG 1501
QY 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 1502 AAGAAG-----GATAGTGGCTTCCAGATGAACAGCTTCGAGGC 1540
QY 142 LysLys 143
Db 1541 AAGAAG 1546
RESULT 32
US-09-285-310-1
; Sequence 1, Application US/09285310
; Patent No. 6262026
; GENERAL INFORMATION:
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; APPLICANT: Concino, Michael F.
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE
; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285,310
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,188
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Ph.D., J.D., Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 07236/009002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3427 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna to mRNA
US-09-285-310-1

Alignment Scores:
Pred. No.: 2.63 Length: 3427
Score: 80.00 Matches: 34
Percent Similarity: 40.16% Conservative: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.36% Indels: 20
DB: 4 Gaps: 6

US-09-995-493-52 (1-179) x US-09-285-310-1 (1-3427)
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 1211 ACTGTGAGATGGTGTGCGAGTCGCGAGCACTAAGTCCAGAGTTTCGCGAC 1270
QY 52 --LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 70
Db 1271 CATATGAAGCGTCATTCATCCAGTGGTCCAGTGTCTGTGTGAAGAAAGCCCTCC 1330
QY 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThr 90
Db 1331 TACCTTGATTCATCAGG---GCCATTGCGCGCAACGAGCG-----GATGCTGTGACA 1381
QY 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSer----- 107
Db 1382 CTGGATCGAGTTGGTGTATGATGCTTACTTGGCTCCCAATAACCTGAACGCTGTGGTG 1441
QY 108 -----AlaSerIleProAspThrAspArgMetTyrTyrSerLeuAsnAspAlaLeuThr 123
Db 1442 GCAGAGTTCTATGGTCAAAAGAGGATCCACAGACTTCTATTATGCTGTGTGGTG 1501
QY 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 1502 AAGAAG-----GATAGTGGCTTCCAGATGAACCAAGCTTCGAGGC 1540

QY 142 LysLys 143
Db 1541 AAGAAG 1546

RESULT 33
US-09-470-058-3
; Sequence 3, Application US/08470058
; Patent No. 5817789
; GENERAL INFORMATION:
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; TITLE OF INVENTION: Chimeric Proteins For Use in Transport
; TITLE OF INVENTION: of a Selected Substance Into Cells
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,058
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3427 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna to mRNA
US-09-285-310-1

Alignment Scores:
Pred. No.: 2.63 Length: 3427
Score: 80.00 Matches: 34
Percent Similarity: 40.16% Conservative: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.36% Indels: 20
DB: 4 Gaps: 6

US-09-995-493-52 (1-179) x US-08-470-058-3 (1-4601)
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 2218 ACTGTGAGATGGTGTGCGAGTCGCGAGCACTAAGTCCAGAGTTTCGCGAC 2277
QY 52 --LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 70
Db 2278 CATATGAAGCGTCATTCATCCAGTGGTCCAGTGTCTGTGTGAAGAAAGCCCTCC 2337
QY 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThr 90
Db 2338 TACCTTGATTCATCAGG---GCCATTGCGCGCAACGAGCG-----GATGCTGTGACA 2388
QY 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSer----- 107
Db 2389 CTGGATCGAGTTGGTGTATGATGCTTACTTGGCTCCCAATAACCTGAACGCTGTGGTG 2448
QY 108 -----AlaSerIleProAspThrAspArgMetTyrTyrSerIleGlyAlaThr 123
Db 2449 GCAGAGTTCTATGGTCAAAAGAGGATCCACAGACTTCTATTATGCTGTGTGGTG 2508
QY 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 2509 AAGAAG-----GATAGTGGCTTCCAGATGAACCAAGCTTCGAGGC 2547

QY 142 LysLys 143
Db 2548 AAGAAG 2553

RESULT 34
US-09-037-188-3
; Sequence 3, Application US/09037188
; Patent No. 6027921
; GENERAL INFORMATION:
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; APPLICANT: Concino, Michael F.
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE
; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,058
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
```


Oy 20 ProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyr 39
Db 632 CCTAAAGGTTTAACTGAAGTACCGTATTATGAATCGATGAATCGGAGCTAATAGCT 691
Oy 40 SerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db 692 TCGTTTGTCTATACATCAT-----LysAspAsnSer 75
Oy 60 SerGlyTyrGluAlaPheThrLysLysGluGluTyr-----LysAspAsnSer 75
Db 710 CAGGATATGATTTCTTCATGCGCATGAATAAGTTTGGTCATGCGATGTTGATTACTACT 769
Oy 76 ArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArgAla----- 93
Db 770 TCAGTAACAATGGGCCCATCTTTCGCATCAACGAATATGTTAGCTTTATGATTACTG 829
Oy 94 GlyLeuAlaTyrAspLys-----AlaAlaSerLys 103
Db 830 GGGCGCGCTCATGGAAGGTTAAGCATCTCTGTTTGTATGATCAATCAGTCAAGCAAG 889
Oy 104 ThrHisLeuSerAlaSerIleProAspThrArgMetTrpTyrSerIleGlyAlaThr 123
Db 890 ACCTCAATGGCA-----TACGGCGCAGGGGTGCA 919
Oy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
Db 920 TTCAACCCACTTCCAAATTTTCTCATTCAGCTTCATATGAATAC 964
RESULT 38
US-08-628-434-1
; Sequence 1, Application US/08628434
; Patent No. 6265567
; GENERAL INFORMATION:
; APPLICANT: Sparling, P. Frederick
; APPLICANT: Beucher, Margaret
; TITLE OF INVENTION: Isolated FrpB Nucleic Acid Molecule and Vaccine
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,434
; FILING DATE: 05-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/418,964
; FILING DATE: 07-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallagher, Thomas C.
; REGISTRATION NUMBER: 37,066
; REFERENCE/DOCKET NUMBER: SPA-3-P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 645-1405
; TELEFAX: (212) 645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 318..2456

US-08-628-434-1
Alignment Scores:
Pred. No.: 3,73 Length: 2600
Score: 77.50 Matches: 41
Percent Similarity: 33.83% Conservative: 27
Best Local Similarity: 20.40% Mismatches: 76
Query Match: 8.10% Indels: 57
DB: 4 Gaps: 10
US-09-995-493-52 (1-179) x US-08-628-434-1 (1-2600)
Oy 7 GlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSer 26
Db 1647 GCACAAACCGTTTCAAGCAGCAGCCTCAACCCGAGTTTCGGCGTGATTGG----- 1697
Oy 27 GlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyr----- 43
Db 1698 -----CAGCCGCGCGCAACACTGGAGCTTCAGCGCGCAGCCACAACATACCGCGCGCG 1748
Oy 43 ----- 43
Db 1749 ACCCGCGCGCTGTATGACGCTCTGCAAAACCCAGCGGCGCATCATCTCGATTGCG 1808
Oy 44 -----ThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGly 59
Db 1809 GACGGCAGCAAGCGCGCGCGCAATACCGAAATCGGCTTCAACTACACGACGCG 1868
Oy 60 SerGlyTyrGluAlaPheThrLysLysGluGluTyr-----LysAspAsnSerArgPheAla 78
Db 1869 ACG-----TTTGGCGCAACCGCGCAGCTACTTCGGCGCAGACCATCAACGACGCG 1916
Oy 79 IleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAsp 98
Db 1917 CTTCGCAATCCCAAAACCGCGCAGCTCCGTCGCGCGCGGAA-----GCCGTCAAC 1970
Oy 99 LysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
Db 1971 GCGCGCTACATCAAAACAC-----GGTTAC 1997
Oy 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
Db 1998 GAATTGGCGCGCTCTCTACCGC---ACCGCGCGCGCTGACCGCCAAAGTCGCGCTAAGCCAC 2054
Oy 139 LeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuVal 158
Db 2055 -----AGCAACCGCGCTTTTACGATACGACCAAGACAAG---CTGTTGACCGCG 2102
Oy 159 GluAlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArg 178
Db 2103 AACCTGAATTTGGCGCACAA---GTCGGCGCGCACTTGGACGCGCTCCCTTCCTACCGC 2159
Oy 179 Phe 179
Db 2160 TTC 2162
RESULT 39
US-09-277-565-24/C
; Sequence 24, Application US/09277565
; Patent No. 6207384
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; APPLICANT: Akerley, Brian J.
; APPLICANT: Rubin, Eric J.
; APPLICANT: Camilli, Andrew
; TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL
; FILE OF INVENTION: GENES BY IN VITRO TRANSPOSON MUTAGENESIS
; FILE REFERENCE: 00742/052002
; CURRENT APPLICATION NUMBER: US/09/277,565
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/079,770
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 29

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-09-277-565-24

Alignment Scores:
Pred. No.: 0.885 Length: 852
Score: 77.00 Matches: 30
Percent Similarity: 45.54% Conservative: 16
Best Local Similarity: 29.70% Mismatches: 47
Query Match: 8.05% Indels: 8
DB: 4 Gaps: 2

US-09-995-493-52 (1-179) x US-09-277-565-24 (1-852)
Qy 73 AspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArg 92
Db 494 GATTCATAAAATCAAGTGCCTCTCTCTCGGATATTTTCTTAATTTATCTTTCCGT 435
Qy 93 AlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAsp 112
Db 434 GCTTTACTTTTCGAGGTAAAGCCTTCAAAACAAATCTATCAGAAGCAATCCCGAT 375
Qy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
Db 374 GCACAAAGAGCGGTAAATGCGCGACAGCTCCTCGCAAGGACACAACTCGAATGCCAGCT 315
Qy 133 AspValGlyPheAlaHisLeuArgGlyLysLysLys-----HisPheVal 147
Db 314 TCA-----CGGCAATGCGGTACTAATGAAMACCAGGTCACGTAAATGCGGTC 264
Qy 148 GluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrLysAlaThr 167
Db 263 CCCGATCAGAAATCAAGGCAATATTACTCCCTCGCTTGAGCTTTCCACCAAAATATGC 204
Qy 168 Ala 168
Db 203 GCT 201

RESULT 40
US-09-087-277-1
; Sequence 1, Application US/09087277B
; Patent No. 6163226
; GENERAL INFORMATION:
; APPLICANT: EK, BO
; APPLICANT: KHOSNOODI, Jamshid
; APPLICANT: LARSSON, Clas-Tomas
; APPLICANT: LARSSON, Hakan
; APPLICANT: RASK, Lars
; TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO
; FILE REFERENCE: 003300-486
; CURRENT APPLICATION NUMBER: US/09/087,277B
; CURRENT FILING DATE: 1998-05-29
; EARLIER APPLICATION NUMBER: PCT/SE96/01558
; EARLIER FILING DATE: 1996-11-28
; EARLIER APPLICATION NUMBER: SE 9504272-7
; EARLIER FILING DATE: 1995-11-29
; EARLIER APPLICATION NUMBER: SE 9601506-0
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3074
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: beII gene
; OTHER INFORMATION: (branching enzyme II) from Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (189)..(2825)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (189)..(332)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (333)..(2825)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (92)..(2156)
; OTHER INFORMATION: Nucleotides 92, 285, 1406, 1430, 1897 and 2156 are
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285)..(287)
; OTHER INFORMATION: Amino acid -16 is Xaa wherein Xaa = Ile, Leu, Val
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1404)..(1406)
; OTHER INFORMATION: Amino acid 358 is Xaa wherein Xaa = Leu or Phe.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1428)..(1430)
; OTHER INFORMATION: Amino acid 366 is Xaa wherein Xaa = Thr.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1896)..(1898)
; OTHER INFORMATION: Amino acid 522 is Xaa wherein Xaa = Tyr, Ser, Cys
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2154)..(2156)
; OTHER INFORMATION: Amino acid 608 is Xaa wherein Xaa = Pro.
US-09-087-277-1

Alignment Scores:
Pred. No.: 7.4 Length: 3074
Score: 76.00 Matches: 16
Percent Similarity: 56.52% Conservative: 10
Best Local Similarity: 34.78% Mismatches: 18
Query Match: 7.94% Indels: 2
DB: 4 Gaps: 1

US-09-995-493-52 (1-179) x US-09-087-277-1 (1-3074)
Qy 38 HistTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
Db 765 CACCTTGATTACAGGTATTTCACAGTACAAAGAACTGAGGAGGCAATTGACAAGTATGAG 824
Qy 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluLysLysAspAsnSerArgPhe 77
Db 825 -----GGTGGTTTGGAAAGCTTTTCTCGTGGTATGAAAAAATGGGTTTTCACCTCGTAGT 878
Qy 78 AlaIleGlyThrThrTyr 83
Db 879 GCTACAGGTATCATTAC 896

Search completed: May 19, 2003, 19:58:38
Job time : 65 secs
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 19, 2003, 19:53:07 ; Search time 99 Seconds
(without alignments)
2332.312 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 OHNGVLGPIYIGKSLTKLP.....ADYTTKATNLYGLNLNRYF 179

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 16200014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+p2n.model
-O=/cgn2_1/USPTC_spool/US0995493/runat_12052003_091311_23115/app_query.fasta_1.327
-DB=PublishedApplications_NA -QPM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	549	57.4	1598	10 US-09-452-599-26	Sequence 26, Appl
2	133.5	13.9	1371	10 US-09-997-664-103	Sequence 103, Appl
3	133.5	13.9	1562	10 US-09-997-664-91	Sequence 91, Appl
4	87	9.1	23451	9 US-10-114-170-173	Sequence 173, Appl

c

5	83.5	8.7	784	10	US-09-765-272-151	Sequence 151, Appl
6	83	8.7	9057	9	US-10-114-170-194	Sequence 194, Appl
7	81.5	8.5	1359	10	US-09-815-242-4086	Sequence 4086, Ap
8	80.5	8.4	4605	9	US-10-080-505-10	Sequence 10, Appl
9	80.5	8.4	4822	9	US-10-080-505-14	Sequence 14, Appl
10	80	8.4	547	9	US-09-854-133-169	Sequence 169, Appl
11	80	8.4	547	10	US-09-738-973-169	Sequence 169, Appl
12	80	8.4	2097	9	US-09-826-025-25	Sequence 25, Appl
13	80	8.4	2347	10	US-09-880-107-3040	Sequence 3040, Ap
14	79.5	8.3	1992	9	US-10-098-841-310	Sequence 310, Appl
15	78	8.2	493	10	US-09-867-550-289	Sequence 289, Appl
16	78	8.2	1204	10	US-09-881-752A-187	Sequence 187, Appl
17	77.5	8.1	554	9	US-10-066-543-2097	Sequence 2097, Ap
18	77	8.0	852	10	US-09-815-242-7171	Sequence 7171, Ap
19	76.5	8.0	9047	10	US-09-070-927A-12	Sequence 12, Appl
20	76	7.9	2160	9	US-09-938-842A-628	Sequence 628, Appl
21	76	7.9	3074	9	US-10-254-534-1	Sequence 1, Appl
22	74	7.7	849	12	US-10-062-994-5	Sequence 5, Appl
23	74	7.7	849	12	US-10-062-994-5	Sequence 5, Appl
24	73.5	7.7	5253	9	US-10-080-943-3	Sequence 3, Appl
25	73.5	7.7	5253	10	US-09-880-107-1655	Sequence 1655, Ap
26	73.5	7.7	7304	9	US-10-114-170-174	Sequence 174, Appl
27	73	7.6	1022	10	US-09-850-351A-33	Sequence 33, Appl
28	73	7.6	15261	9	US-09-764-891-10186	Sequence 10186, A
29	73	7.6	124884	9	US-09-913-514-1	Sequence 1, Appl
30	73	7.6	125157	9	US-09-913-514-2	Sequence 2, Appl
31	72.5	7.6	1254	10	US-09-925-297-58	Sequence 58, Appl
32	72.5	7.6	1380	10	US-09-969-347-209	Sequence 209, Appl
33	72	7.5	1239	10	US-09-974-300-4357	Sequence 4357, Ap
34	72	7.5	4305	9	US-10-080-505-8	Sequence 8, Appl
35	71.5	7.5	807	9	US-09-938-842A-212	Sequence 212, Appl
36	71	7.4	636	10	US-09-974-300-1094	Sequence 1094, Ap
37	71	7.4	1548	9	US-10-114-170-86	Sequence 86, Appl
38	70.5	7.4	1365	10	US-09-822-849A-513	Sequence 513, Appl
39	70.5	7.4	1464	9	US-09-738-626-216	Sequence 216, Appl
40	70.5	7.4	1585	10	US-09-765-205-11	Sequence 11, Appl
41	70.5	7.4	1631	10	US-09-864-864-301	Sequence 301, Appl
42	70.5	7.4	1638	9	US-09-905-291A-260	Sequence 260, Appl
43	70.5	7.4	1638	9	US-09-902-853-260	Sequence 260, Appl
44	70.5	7.4	1638	9	US-09-907-824-260	Sequence 260, Appl
45	70.5	7.4	1638	9	US-09-907-841-260	Sequence 260, Appl

ALIGNMENTS

RESULT 1
US-09-452-599-26
; Sequence 26, Application US/09452599
; Patent No. US20020055101A1
; GENERAL INFORMATION:
; APPLICANT: Bergeron, Michel G.
; APPLICANT: Ouellette, Marc
; APPLICANT: Roy, Paul H.
; TITLE OF INVENTION: Specific and Universal Probes and Amplification Primers
; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical
; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
; FILE REFERENCE: 12287.31
; CURRENT APPLICATION NUMBER: US/09/452,599
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 08/526,840
; PRIOR FILING DATE: 1995-09-11
; PRIOR APPLICATION NUMBER: 08/304,732
; PRIOR FILING DATE: 1994-09-12
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-09-452-599-26
Alignment Scores:

Pred. No.: 8,54e-65 Length: 1598
Score: 549.00 Matches: 105
Percent Similarity: 76.74% Conservatve: 27
Best Local Similarity: 61.05% Mismatches: 34
Query Match: 57.37% Indels: 6
DB: 10 Gaps: 2

US-09-995-493-52 (1-179) x US-09-452-599-26 (1-1598)

Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeu 31
Db 939 AAAGTAATTACCTTTACATTGCCAGATTCTGAGATCTTCTGGTTTCCATCAATTA 998
Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 999 ACTGACAACTTCGAGTGCATATAGTTATATAATATACCCATCGGCTGCTTTAACAAAA 1058
Qy 52 LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyr 71
Db 1059 TTACATGCCAGCTTCGAAGATGGTAAA-----AAAGCTTTTGATAAAGAAATTACAATAC 1112
Qy 72 LysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeu 91
Db 1113 AGTAATACCTCTCGTGTGATTAGGGCGCAAGTTATATCTTTATGAAAAATTGACCTTA 1172
Qy 92 ArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIlePro 111
Db 1173 CGTGGGGTATGCTTACGATCAAGCGGCATCTCGTCATCACCGTAGTGTGCAATTTCCA 1232
Qy 112 AspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 1233 GATACCATCGCACTTGGTATAGTTAGTGGCAACCTTATAAATTCACGCCGAATTTATCT 1292
Qy 132 ValAspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsn 151
Db 1293 GTTGATCTTGGCTATGCTTACTTAAAGGCAAAAAGTTTACCTTTAAAGAAAGTAAAAACA 1352
Qy 152 IleLysGly-----LeuLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 1353 ATAGGTGACAAACGTACATTGACATTGAATACAACTGCAAAATTATCTTCAAGCACAC 1412
Qy 168 AlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 1413 GCAATCTTTACGTTTGAATTTAAATATATAGTTTC 1448

RESULT 2

US-09-997-664-103
; Sequence 103, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Cattermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sarlasani, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; FILE REFERENCE: BC1018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/585,174
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 103
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Pseudomonas mendocina KR-1

US-09-997-664-103

Alignment Scores:

Pred. No.: 3.45e-08 Length: 1562
Score: 133.50 Matches: 40
Percent Similarity: 44.00% Conservatve: 26
Best Local Similarity: 26.67% Mismatches: 75

Pred. No.: 2.85e-08 Length: 1371
Score: 133.50 Matches: 40
Percent Similarity: 44.00% Conservatve: 26
Best Local Similarity: 26.67% Mismatches: 75
Query Match: 13.95% Indels: 9
DB: 10 Gaps: 4

US-09-995-493-52 (1-179) x US-09-997-664-103 (1-1371)

Qy 13 GlySerLeuThrLeuLys-----LeuProAlaTyrTrpGluLeuSerGlyPheHis 29
Db 871 GGGGATATCCGCGTAAAGACTTCGAGATGCCCGCCAGCTGACGTTGGCTTTCATCAT 930
Qy 30 GlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg--- 48
Db 931 CAATTCAACGAGCGTTCGCTGCTGCTCATGCTCAAGCGTCTCTACTGGAGCGATGTC 990
Qy 49 PheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
Db 991 ATGAGACACATCAGTGTGGATTTCAAAATCGCAGTCAGGTGGGATTGATATCGAATTACCA 1050
Qy 69 GluLutyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAla 88
Db 1051 CACAACATCAGGATATACGGTGGCTCCATCGCACCCTTACAGAGTTATGACAAG 1110
Qy 89 LeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAla 108
Db 1111 CTAACTCTTCGCTGCTGATATAGCTATCGCAACGCGCTGGACAGTAGCTGATATTG 1170
Qy 109 Ser---IleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThr 127
Db 1171 CCAGTAATTTCCAGCTTATTTTGAAGAAACACAGTTTCTCGGTAGCGATTATCTTTTGTAT 1230
Qy 128 Pro-----AsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLys 143
Db 1231 AAAAATCAAACTCAATTTGGCGATTCTTTTGGCCCTAAAGAGACCTTGAACACACCA 1290
Qy 144 LysHisPheValGluThrGlnAsnIleLys 153
Db 1291 TCATACCTAAGCGCGCACCGAAGCTTGAAG 1320

RESULT 3

US-09-997-664-91
; Sequence 91, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Cattermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sarlasani, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; FILE REFERENCE: BC1018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/585,174
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 91
; LENGTH: 1562
; TYPE: DNA
; ORGANISM: Pseudomonas mendocina KR-1

US-09-997-664-91

Alignment Scores:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765.272.
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961.083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 151:
SEQUENCE CHARACTERISTICS:
LENGTH: 784 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 151:
US-09-765-272-151
Alignment Scores:
Pred. No.: 0.0819 Length: 784
Score: 83.50 Matches: 42
Percent Similarity: 39.53% Conservative: 26
Best Local Similarity: 24.42% Mismatches: 77
Query Match: 8.73% Indels: 28
DB: 10 Gaps: 7
US-09-995-493-52 (1-179) x US-09-765-272-151 (1-784)
QY 8 ProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGly 27
Db 23 CCGTTTCATTAAGGTAGTACAAATGGTCCCTTACCCTAACCAATGGTATCAGATTAGAAGGT 82
QY 28 PheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLys-----TyrThr 44
Db 83 TTTCCAAG-----TCAGAGTGG-----TACTACTTCGATAAAATGGAGTCTACAA 130
QY 45 GluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAla 64
Db 131 GAGTTGTTGGTTGGAAACATTAGAGATTAAACTAAAGACAGTGTGGGAAGAAAGTAC 190
QY 65 PheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrSer 84
Db 191 GGGGAAACCGTGAAGATTCAGAGATTAAGAGAGAGCGCTTATTATACCAACTATTAC 250
QY 85 LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla----- 100
Db 251 TTTAATCAAAATCATTTTAGACACAGGTGGCTTTATGATCAGTCTAAGTGTATTAT 310
QY 101 AlaSerLysThrHisLeuSer-----AlaSerIle 110
Db 311 CTAGCTAAGACGGAATTAATGGAGAAACACTCCTTGGTGTGAAAGACGTCGGGGTGG 370
QY 111 ProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu 130

Db 371 ATAAACGATGATTCCGACTTGGTACTACCTAGATCCAAACAACCTGCTATT----- 418
QY 131 SerValaspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThrCln 150
Db 419 ---ATGCAACAGGTGGCAATATCTA---GGTAATAAGTGGTACTACCTCCGTTCTCTCA 472
QY 151 AsnIleLysGlyLeuLeuValGluAlaAspTyr 162
Db 473 GGAGC-AATGCCCACTGGCTGTATCAGGAAGGTAC 507
RESULT 6
US-10-114-170-194/c
Sequence 194, Application US/10114170
Publication No. US20030023075A1
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114.170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453.702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110.955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 194:
SEQUENCE CHARACTERISTICS:
LENGTH: 9057
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 194:
US-10-114-170-194
Alignment Scores:
Pred. No.: 3.45 Length: 9057
Score: 83.00 Matches: 42
Percent Similarity: 37.02% Conservative: 25
Best Local Similarity: 23.20% Mismatches: 58
Query Match: 8.67% Indels: 56
DB: 9 Gaps: 10
US-09-995-493-52 (1-179) x US-10-114-170-194 (1-9057)
QY 33 AspGlnTrpAlaIleHis-----TyrSerTyrLysTyrThrGluTrpSerArgPheLys 50
|||||:|||||:|

```
Db 1897 GACAAATGGTCATCTCGTGGGGGATGACTATCAATCCGACTAACTGGGCTGATG----- 1844
QY 51 GluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
   ||| ||| |||
   : : : : : ||| |||
Db 1843 ---TTATTTGGCTCATAT-----GCCAGGCATTCGGCGCCCGACGATGGCGMAATG 1793
QY 71 TyrLysAspAsnSerArgPheAlaIleGlyThrTyrSer----- 84
   ||| ||| : : : : : ||| |||
   : : : : : ||| |||
Db 1792 TATAACGATTCTAAGCACTTCTCGATTGGTCGCTTCTATACCAACTATTGGGTGCCAAAC 1733
QY 84 ----- 84
Db 1732 CCGAACTTACGTCGCGGAAACTACGAAACTCAGAGTACGGTTTGGGCTGCGCTTTTGAT 1673
QY 85 -----LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
   ||| ||| ||| ||| ||| |||
   : : : : : : : : : : :
Db 1672 GACCTGATGTTGTCCAATGATCTCTGGAATTTAAAGCCAGCTACTTTTGATACCAAGCG 1613
QY 101 -----AlaSerLysThrHisLeuSerAlaSerIle 110
   ||| : : : : : ||| ||| |||
   : : : : : ||| ||| |||
Db 1612 AAGGATTACATCTCCACGACCGTCGATTTGCGCGCGCGGACGACTATGTCGTATAACGTC 1553
QY 111 ProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeu 130
   ||| : : : : : ||| ||| |||
   : : : : : ||| ||| |||
Db 1552 CCGAACGCC---AAAATCTGG---GGCTGGGATGTGATCAGCAATATATACCACTGATCTG 1499
QY 131 ---SerValAspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThr 149
   ||| ||| ||| ||| ||| |||
   : : : : : : : : : : :
Db 1498 TTTAGCCTTGATGTGGCTATACCGTACCGCGCGGANA-----GACACC 1454
QY 150 GlnAsnIleGlyLysLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsn 169
   : : : : : ||| ||| ||| |||
   : : : : : ||| ||| |||
Db 1453 GATACCGCGGAATACATCTCCAGCATTAACCCGGAT---ACTGTACCAAGCACTCTGAAT 1397
QY 170 Leu 170
Db 1396 ATT 1394
```

RESULT 7

```
US-09-815-242-4086
; Sequence 4086, Application US/09815242
; Patent No. US20020061569A1
; ORGANISM: Pseudomonas aeruginosa
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wali, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4086
```

```
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4086

Alignment Scores:
Pred. No.: 0.343 Length: 1359
Score: 81.50 Matches: 35
Percent Similarity: 40.67% Conservative: 26
Best Local Similarity: 23.33% Mismatches: 56
Query Match: 8.52% Indels: 33
DB: 10 Gaps: 7

US-09-995-493-52 (1-179) x US-09-815-242-4086 (1-1359)
QY 4 GlyValLeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyr--- 22
   ||| : : : ||| ||| ||| ||| ||| |||
   : : : : : ||| ||| ||| ||| |||
Db 703 GGCTAGCTCGCGCCCTACACCTTCGCCCCAGCACCTGCAGCAACGCCCTGCCGCGCTACGAA 762
QY 23 -----TrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSer 40
   ||| : : : ||| ||| ||| ||| |||
   : : : : : ||| ||| ||| ||| |||
Db 763 AGCCACTTCAAGCAGACGCGC---AAGCAGCTGATACCGACTGGCGCCTGCTCGCGCGCC 819
QY 41 TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySer 60
   ||| ||| ||| ||| ||| ||| |||
   : : : : : ||| ||| ||| ||| |||
Db 820 ATCGGCTTACCAGGAA-----TCGCTGTGGCAGCGCGCGCC 855
QY 61 GlyTyrGluAlaPheThrLysLysGluGluTyrLys-----AspAsn 74
   ||| ||| ||| ||| ||| |||
   : : : : : ||| ||| ||| ||| |||
Db 856 -----ACCTCCAAGACCGCGCTGGCGGCTGATGATGTCGACCAACCGG 900
QY 75 SerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGly 94
   : : : ||| : : : ||| ||| ||| |||
   : : : ||| : : : ||| ||| ||| |||
Db 901 ACCGCCAGCGGATGGCGTGTCCACCGCGCTCGACCCGAGCAGACATCCAGGCGCGC 960
QY 95 LeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAsp 114
   ||| ||| ||| ||| ||| ||| |||
   : : : ||| ||| ||| ||| ||| |||
Db 961 AGCAAGTAT---TTCGTGCAGATCCCGACGCAACTGCCCCGAGAGCATCAAGAACCGCAC 1017
QY 115 ArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspVal 134
   ||| ||| ||| ||| ||| ||| |||
   : : : ||| ||| ||| ||| ||| |||
Db 1018 CGCAGCTGGTTCGCCCTGGCGCGC-----TACAACATC 1050
QY 135 GlyPheAlaHisLeuArgGlyLysLysLys 144
   ||| ||| ||| ||| ||| |||
   : : : ||| ||| ||| ||| ||| |||
Db 1051 GCGCGCGCGCACCTGGGAAGACGCGCGCAAG 1080

RESULT 8
US-10-080-505-10
; Sequence 10, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: patentIn version 3.1
; SEQ ID NO 10
; LENGTH: 4605
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (422)..(4597)
; OTHER INFORMATION:
US-10-080-505-10
```

Alignment Scores:
 Pred. No.: 2.8 Length: 4605
 Score: 80.50 Matches: 34
 Percent Similarity: 38.99% Conservative: 28
 Best Local Similarity: 21.38% Mismatches: 51
 Query Match: 8.41% Indels: 46
 DB: Gaps: 7

US-09-995-493-52 (1-179) x US-10-080-505-10 (1-4605)

QY 22 TyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyr 41
 Db 994 CATGTGAGACTTACCTATGAGGTTTCATGGTTAAATAGTGG-----AAATACCTTTGA 1047
 QY 42 LysTyrThrGluTrpSerArgPhe-LysGluLeuArgGlyLysTyrGlnAspGlySerG1 61
 Db 1048 AGATGACCAGCTGGTAAACGGTACATTAGAAATTAATGGCGAGTACAAAATCCTAATAA 1107
 QY 61 yTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyTh 81
 Db 1108 ATATGCTCCACTACCTACGCGAGGTTTCATTCGGGGATAGT-----GGTTC 1152
 QY 81 rThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 101
 Db 1153 TCCAATGTTTATT-----TATCATAAGGAAGT 1179
 QY 101 aserLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleG1 121
 Db 1180 TAAGAAATGGTTTAAATGGCGGTGTTACCTGAA---GGAAATCCTTATCCTCCAGTAGG 1236
 QY 121 yAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgG1 141
 Db 1237 AACAGCTATCAAAATTACA----- 1255
 QY 141 yLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAs 161
 Db 1256 -CGAAAGATTTATTT-----CAAGGTATTTCTT----- 1282
 QY 161 pTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1283 ----AATCAAGACATTACAGCTAAATTTTGGGATACATAATGCTGAATATAGATTT 1333

RESULT 9
 US-10-080-505-14
 ; Sequence 14, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Geme, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 14
 ; LENGTH: 4822
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (388)..(4563)
 ; OTHER INFORMATION:
 US-10-080-505-14

Alignment Scores:
 Pred. No.: 3 Length: 4822
 Score: 80.50 Matches: 34
 Percent Similarity: 38.99% Conservative: 28
 Best Local Similarity: 21.38% Mismatches: 51

Query Match: 8.41% Indels: 46
 DB: Gaps: 7
 US-09-995-493-52 (1-179) x US-10-080-505-14 (1-4822)
 QY 22 TyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyr 41
 Db 960 CATGTGAGACTTACCTATGAGGTTTCATGGTTAAATAGTGG-----AAATACCTTTGA 1013
 QY 42 LysTyrThrGluTrpSerArgPhe-LysGluLeuArgGlyLysTyrGlnAspGlySerG1 61
 Db 1014 AGATGACCAGCTGGTAAACGGTACATTAGAAATTAATGGCGAGTACAAAATCCTAATAA 1073
 QY 61 yTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyTh 81
 Db 1074 ATATGCTCCACTACCTACGCGAGGTTTCATTCGGGGATAGT-----GGTTC 1118
 QY 81 rThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 101
 Db 1119 TCCAATGTTTATT-----TATCATAAGGAAGT 1145
 QY 101 aserLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleG1 121
 Db 1146 TAAGAAATGGTTTAAATGGCGGTGTTACCTGAA---GGAAATCCTTATCCTCCAGTAGG 1202
 QY 121 yAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgG1 141
 Db 1203 AACAGCTATCAAAATTACA----- 1221
 QY 141 yLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAs 161
 Db 1222 -CGAAAGATTTATTT-----CAAGGTATTTCTT----- 1248
 QY 161 pTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
 Db 1249 ----AATCAAGACATTACAGCTAAATTTTGGGATACATAATGCTGAATATAGATTT 1299

RESULT 10
 US-09-854-133-169
 ; Sequence 169, Application US/09854133
 ; Publication No. US20020183499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.475C10
 ; CURRENT APPLICATION NUMBER: US/09/854,133
 ; CURRENT FILING DATE: 2001-05-11
 ; NUMBER OF SEQ ID NOS: 735
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 169
 ; LENGTH: 547
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-854-133-169

Alignment Scores:
 Pred. No.: 0.145 Length: 547
 Score: 80.00 Matches: 34
 Percent Similarity: 40.16% Conservative: 15
 Best Local Similarity: 27.87% Mismatches: 53
 Query Match: 8.36% Indels: 20
 DB: Gaps: 6

US-09-995-493-52 (1-179) x US-09-854-133-169 (1-547)
 QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
 Db 124 ACTGTGATGATGGTGTGCAGTGTCCGAGCATGAGCCACCAATGATGCCAGACTTCCCGGAC 183

Query Match: 8.31% Indels: 31
DB: 9 Gaps: 6

US-09-995-493-52 (1-179) x US-10-098-841-310 (1-1992)

QY 30 GlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPhe 49
DB 965 AAAATTCGAGACTATGCTTTTGGCACTTCAGTAACCGAGAAGATGCAGTTGAGGCTATG 1024
QY 50 LysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
DB 1025 AAAGCTTTAAATGCGAAGGTGCTGATGGTTCCCAATGGAAGTCAACCTAGCAAAACCA 1084
QY 70 GluTyrLysAspAsn---SerArgPheAlaIleGlyThr----- 81
DB 1085 GTGGACAAGGACAGATTATGTTAGTATACCCGAGGCACAGGTGGAAGGGCCACCATGCTG 1144
QY 82 -----ThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
DB 1145 CAAGGAGACTATACCTACTCTTTG-----GGCCAAGTTTAT 1180
QY 98 AspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrp 117
DB 1181 GAT-----CCCACCACACCTACCTTGGAGCTCCTGTC----- 1213
QY 118 TyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAla 137
DB 1214 TTCTATGCCCCCAGACCTATGACGCAATTCGCCAGTCTTCATTTCCAGCCACCAAGGA 1273
QY 138 HisLeuArgGlyLysLysHisPheValGluThrGlnAsnIleLysGly 154
DB 1274 CATCTC-----AGCAACAGAGCCATTATCCGAGCCCTTCTGTAGAGG 1318

RESULT 15

US-09-867-550-289

; Sequence 289, Application US/09867550

; Patent No. US20020082206A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Mehraban, Fuad,

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

; FILE REFERENCE: 21402-013 (Cura-313)

; CURRENT APPLICATION NUMBER: US/09/867,550

; CURRENT FILING DATE: 2001-09-20

; PRIOR APPLICATION NUMBER: USSN 60/208,427

; PRIOR FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 2125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 289

; LENGTH: 493

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)

; OTHER INFORMATION: wherein any n ls one of a or t or g or c

US-09-867-550-289

Alignment Scores:

Pred. No.: 0.233 Length: 493

Score: 78.00 Matches: 37

Percent Similarity: 36.55% Conservatlve: 16

Best Local Similarity: 25.52% Mismatches: 56

Query Match: 8.15% Indels: 36

DB: 10 Gaps: 6

US-09-995-493-52 (1-179) x US-09-867-550-289 (1-493)

QY 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla 36

DB 33 CTGAACACCTCCGCGGACTGCGGCTGTCTCAGGTGGAACACCTGTCTTCCGCTGG--- 89
QY 37 IleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyr 56
DB 90 -----TGGTAGCGCTACACGGTGCATGAGGAACTGTGCGACTTCCGCGCGCTCACC 140
QY 57 GlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu----- 69
DB 141 TCGGACGGCAACCACTATCTCGGCTACAGCTCGCAGGTGATCGCGTGGGGCAAGATC 200
QY 70 -----GluTyrLysAspAsnSerArgPheAlaIleGlyThr 81
DB 201 CTCTCGGCTACGACCTCAACTTCATGCAGCAGGTCAACACGCTGTTTCATCGCGACACC 260
QY 82 ThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAla 101
DB 261 CAGCGCTGCTCAACGACAGGCTCACCCCTCAGCGCGGGCTTC-----CGCGCGCC 311
QY 102 SerLysThrHisLeuSerAlaSer---IleProAspThrAspArgMetTrpTyrSerIle 120
DB 312 ATGCTCTCGCGCCGAGGCCACCACTCATCCCGCGCGCGATCCC-----TACAAGATG 365
QY 121 -----GlyAlaThrTyrLysPheThrProAsn 129
DB 366 CAGGCAACTACTTCGAGCGGCTGCGCAGTTCCTCGGCCAGCTACCAAGATCACCCCGCAC 425
QY 130 LeuSerValAspVal 134
DB 426 GATCAGGTGTGTCGTC 440

RESULT 16

US-09-881-752A-187

; Sequence 187, Application US/09881752A

; Patent No. US20020115078A1

; GENERAL INFORMATION:

; APPLICANT: Kleenathous, Harold

; APPLICANT: Al-Garawi, Amal

; APPLICANT: Miller, Charles

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Oomen, Raymond P.

; TITLE OF INVENTION: Identification of Polynucleotides

; TITLE OF INVENTION: Encoding No. US20020115078A1e1 Helicobacter Polypeptides in

; FILE REFERENCE: 06132/041002

; CURRENT APPLICATION NUMBER: US/09/881,752A

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 08/833,457

; PRIOR FILING DATE: 1997-04-01

; NUMBER OF SEQ ID NOS: 370

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 187

; LENGTH: 1204

; TYPE: DNA

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (51)....(1151)

US-09-881-752A-187

Alignment Scores:

Pred. No.: 0.862 Length: 1204

Score: 78.00 Matches: 35

Percent Similarity: 42.42% Conservatlve: 21

Best Local Similarity: 26.52% Mismatches: 46

Query Match: 8.15% Indels: 30

DB: 10 Gaps: 7

US-09-995-493-52 (1-179) x US-09-881-752A-187 (1-1204)

QY 63 GluAlaPheThrLysLysGluTyrLysAspAsnSerArg---PheAlaIleGlyThr 81
DB 456 GAACCCATTACTAACCCCTTTAGAAATTTGAAGAAACTTAAGAAATTTAGAAAGTCAATTT 515

Db 374 GCACAAAGACGGTAATTCGCCACAAAGCTCTGGCAAAAGGCACAACTCGAATGCCAGCT 315
Qy 133 AspValGlyPheAlaHisLeuArgGlyLysLysLys-----HisPheVal 147
Db 314 TCA-----CGGATCGGCTACTAANTGAACACAGGCTCACTGATTAAATGCCGTC 264
Qy 148 GluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThr 167
Db 263 CCCGCATCAGAAATCAAGCAATATTACTCCCTGCTTGCAGCTTTTCCACCAAAATATGC 204
Qy 168 Ala 168
Db 203 GCT 201
RESULT 19
US-09-070-927A-12
; Sequence 12, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 992
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-070-927A-12
Alignment Scores:
Pred. No.: 26.4 Length: 9047
Score: 76.50 Matches: 41
Percent Similarity: 37.09% Conservative: 15
Best Local Similarity: 27.15% Mismatches: 45
Query Match: 7.99% Indels: 50
DB: 10 Gaps: 6
US-09-995-493-52 (1-179) x US-09-070-927A-12 (1-9047)

Qy 18 LysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIle 37
Db 3630 AAAATTCAGCC-----AAGTTAAGTTTCTTATTAGCGATTACAGAC----- 3671
Qy 38 HisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
Db 3672 -----AGTCGTTTGCCTGTTATGATTCT 3695
Qy 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
Db 3696 GATAATGCTGGGTTTGCATGCTTTTCAAAGAGGCGCACAAAGGCCTTATCAAGACTTT 3755
Qy 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGly-----Leu 95
Db 3756 TTAACAGGC-----AAATTGACCTTCCCACTGTCACACAGCTA 3794
Qy 96 AlaTyrAspLysAlaAlaSerLysThrHisLeuSer----- 107
Db 3795 GCAGAGGATAGTTAGCTCCTAAAATTGTTTATCGAATGTTTTCGAAGTCAACACATGAG 3854
Qy 108 -----AlaSerIleProAspThrAspArgMet 116
Db 3855 TCGATTGAGTCCACGACAGACGGGTCAAGCCCTAAACTACACAGAAAAACAACATATT 3914
Qy 117 TrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGly-Ph 136
Db 3915 TTTTATCGGATGGTTCACACGAGAGTGACTGTCAATTGGTCA-----GAAGTT 3965
Qy 136 eAlaHisLeuArgGlyLysLysLysHisPhe 146
Db 3966 CCAGCATCTTCTTATCAAAAAGAGGAATAT 3996
RESULT 20
US-09-938-842A-628
; Sequence 628, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 628
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-628
Alignment Scores:
Pred. No.: 3.8 Length: 2160
Score: 76.00 Matches: 52
Percent Similarity: 38.43% Conservative: 31
Best Local Similarity: 24.07% Mismatches: 73
Query Match: 7.94% Indels: 60
DB: 9 Gaps: 1
US-09-995-493-52 (1-179) x US-09-938-842A-628 (1-2160)
Qy 3 AsnGly-----ValLeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuPro 20
Db 151 AATGGAGCCACCGTTCTTCTTCGCTATGTTGGTCGATCAGCG-----AAAAAGCTTCCT 204


```

; APPLICANT: Whitlmore, William M.
; APPLICANT: Allemen, Arthur R.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases and
; FILE REFERENCE: UF-167XC3
; CURRENT APPLICATION NUMBER: US/10/062,994
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/533,662
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/337,827
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/953,326
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 08/733,230
; PRIOR FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Anaplasma marginale
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(846)
; US-10-062-994-5

Alignment Scores:
Pred. No.: 1.81 Length: 849
Score: 74.00 Matches: 49
Percent Similarity: 38.00% Conservativeness: 27
Best Local Similarity: 24.50% Mismatches: 86
Query Match: 7.73% Indels: 38
DB: 12 Gaps: 11

US-09-995-493-52 (1-179) x US-10-062-994-5 (1-849)

QY 14 SerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPhe----- 28
Db 250 ACCGTTGATGTGAGTGTGCGAGCAAACTTTCCAAATCTGGCTACACTTTTGCTTCTCT 309
QY 29 ---HisGlnLeuThrAsp---GlnTrpAlaIleHisTyrSerTyr----- 41
Db 310 AAAAAGCTTAATCACTGCTTTTCGACGGCGCTGGGATATTCTCTGGGAGGAGCCAGAGTG 369
QY 42 ---LysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySer 60
Db 370 GAATTCGAAGCGAGCTACAGAGGTTTGCTACTTTGGCGAGCGGAGTACGCAAAAGT 429
QY 61 GlyTyrGlu-----AlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
Db 430 GGTGCGGAATCTCTGGCAGCTATTACCGCGAGCTAACATTACTGAGACCAATTACTTC 489
QY 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
Db 490 GTAGTCAAAATGATGAATCAACAACCTCATGATGTTAAATGGC-----TGCATAT 543
QY 98 AspLysAlaAlaSerLysThrHisLeuSer-----AlaSer. 109
Db 544 GAGCTGCTGCACACAGATTTACCTGTCCTCCCGTATGTATGTCGGGATAGCGCAGC 603
QY 110 IleProAsp-----ThrAspArgMetTrpTyr-----SerIleGlyAla 122
Db 604 TTGTTGACATCTCTAAGCAAGTAACACACAAAGCTGGCTACAGGGGCAAGGTTGGATT 663
QY 123 ThrTyrLysPheThrProAsnLeuSer---ValAspValGlyPheAlaHis-----Leu 139
Db 664 AGCTACAGCTTACTCCGAAATATCTTGGTGGCAGGTGGGTCTTACCACGGGCTATTT 723
QY 140 ArgGlyLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuLeuValGlu 159
Db 724 GATGAGCTTACAGGACATTCCTCCGACACACACAGTGTAAAG-----TTCTCTGGAGAA 777
QY 160 AlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
```

```

Db 778 GCAAAGCCTCAGTCAAGGCGCATATTGCTGACTAGGCTTTAAACCTTGGAGCAAGATTTC 837
RESULT 23
US-10-062-994-5
; Sequence 5, Application US/10062994
; Patent No. US20020132789A1
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Bowler, Michael V.
; APPLICANT: Burridge, Michael J.
; APPLICANT: Mahan, Suman M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Rurangirwa, Fred R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Whitlmore, William M.
; APPLICANT: Allemen, Arthur R.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases and
; FILE REFERENCE: UF-167XC3
; CURRENT APPLICATION NUMBER: US/10/062,994
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/533,662
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/337,827
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/953,326
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 08/733,230
; PRIOR FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Anaplasma marginale
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(846)
; US-10-062-994-5

Alignment Scores:
Pred. No.: 1.81 Length: 849
Score: 74.00 Matches: 49
Percent Similarity: 38.00% Conservativeness: 27
Best Local Similarity: 24.50% Mismatches: 86
Query Match: 7.73% Indels: 38
DB: 12 Gaps: 11

US-09-995-493-52 (1-179) x US-10-062-994-5 (1-849)

QY 14 SerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPhe----- 28
Db 250 ACCGTTGATGTGAGTGTGCGAGCAAACTTTCCAAATCTGGCTACACTTTTGCTTCTCT 309
QY 29 ---HisGlnLeuThrAsp---GlnTrpAlaIleHisTyrSerTyr----- 41
Db 310 AAAAAGCTTAATCACTGCTTTTCGACGGCGCTGGGATATTCTCTGGGAGGAGCCAGAGTG 369
QY 42 ---LysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySer 60
Db 370 GAATTCGAAGCGAGCTACAGAGGTTTGCTACTTTGGCGAGCGGAGTACGCAAAAGT 429
QY 61 GlyTyrGlu-----AlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
Db 430 GGTGCGGAATCTCTGGCAGCTATTACCGCGAGCTAACATTACTGAGACCAATTACTTC 489
QY 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
Db 490 GTAGTCAAAATGATGAATCAACAACCTCATGATGTTAAATGGC-----TGCATAT 543
QY 98 AspLysAlaAlaSerLysThrHisLeuSer-----AlaSer 109
```

```
Db 544 GACGTGCTGCACACAGATTTACCTGCTCCCGTATGATGTCGCCGGATAGCGCAAGC 603
Qy 110 IleProasp-----ThrAspArgMetTrpTyr-----SerIleGlyAla 122
Db 604 TTTGTTGACATCTTAGCAAGTACCAACAAAGCTGCCCTACAGGGCAAGGTGGGATT 663
Qy 123 ThrTyrLysPheThrProAsnLeuSer---ValAspValGlyPheAlaHis-----Leu 139
Db 664 AGTACCAGTTTACTCCGGAATATCCTTGTGCGCAGTGGTCTTACCACGGCTATTT 723
Qy 140 ArgGlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGlu 159
Db 724 GATGAGCTTACCAAGGACATTCGCGCACACACAGTGTAAAG-----TTCTCTGGAGAA 777
Qy 160 AlaAspTyrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyrArgPhe 179
Db 778 GCAAAAGCCTCAGTCAAGGCGCATATTGCTGACTACGGCTTTAACCTTGGAGCAGATTC 837
RESULT 24
US-10-080-943-3
; Sequence 3, Application US/10080943
; Publication No. US20030073236A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; TITLE OF INVENTION: P193 PROTEINS AND NUCLEIC ACIDS, AND USES THEREOF
; FILE REFERENCE: IU99-PCT
; CURRENT APPLICATION NUMBER: US/10/080,943
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/150,266
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(5183)
US-10-080-943-3
Alignment Scores:
Pred. No.: 30.6 Length: 5253
Score: 73.50 Matches: 28
Percent Similarity: 38.46% Conservative: 17
Best Local Similarity: 23.93% Mismatches: 47
Query Match: 7.68% Indels: 25
DB: 9 Gaps: 3
US-09-995-493-52 (1-179) x US-10-080-943-3 (1-5253)
Qy 64 AlapheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyr 83
Db 1134 TCCTTCAGGAGGTCAAGACGTTTTCGCCCTCGTCTGAGTTCGCAAGTGCATACCTAT 1193
Qy 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLys 103
Db 1194 GCITTTGATGTCGGGACACACTGCACCCGGGATGCGAGTGGCGGATGATTAT 1253
Qy 104 ThrHisLeuSerAla-----SerIleProAsp 112
Db 1254 GAGGAGATCATGTCGCCGGGATGAGGGCGAGTTTCGGCAGACAAACACGGTGTGCCTCT 1313
Qy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
Db 1314 GTGCAGGTATTTGGGAGTCAACAGCCGCCATTTATTTGGGTGCACCTGGCATCTCGGAG 1373
Qy 133 AspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIle 152
Db 1374 ATCTTGGGCTTT-----GAGGAAGACATT 1397
US-10-114-170-174
; Sequence 174, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
```

```
Qy 153 LysGlyLeuLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsn 169
Db 1398 GAGGAC-----ATGGTTGAGGCTGATGAGTACCAAGGGCGAGTGGCCAGT 1442
RESULT 25
US-09-880-107-1655
; Sequence 1655, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles In Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1655
; LENGTH: 5253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D38548
US-09-880-107-1655
Alignment Scores:
Pred. No.: 30.6 Length: 5253
Score: 73.50 Matches: 28
Percent Similarity: 38.46% Conservative: 17
Best Local Similarity: 23.93% Mismatches: 47
Query Match: 7.68% Indels: 25
DB: 10 Gaps: 3
US-09-995-493-52 (1-179) x US-09-880-107-1655 (1-5253)
Qy 64 AlapheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyr 83
Db 1134 TCCTTCAGGAGGTCAAGACGTTTTCGCCCTCGTCTGAGTTCGCAAGTGCATACCTAT 1193
Qy 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLys 103
Db 1194 GCITTTGATGTCGGGACACACTGCACCCGGGATGCGAGTGGCGGATGATTAT 1253
Qy 104 ThrHisLeuSerAla-----SerIleProAsp 112
Db 1254 GAGGAGATCATGTCGCCGGGATGAGGGCGAGTTTCGGCAGACAAACACGGTGTGCCTCT 1313
Qy 113 ThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerVal 132
Db 1314 GTGCAGGTATTTGGGAGTCAACAGCCGCCATTTATTTGGGTGCACCTGGCATCTCGGAG 1373
Qy 133 AspValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGluThrGlnAsnIle 152
Db 1374 ATCTTGGGCTTT-----GAGGAAGACATT 1397
US-10-114-170-174
; Sequence 174, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
```

Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1e1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Plinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114, I70
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 174:
SEQUENCE CHARACTERISTICS:
LENGTH: 7304
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 174:
US-10-114-170-174
Alignment Scores:
Pred. No.: 49.6 Length: 7304
Score: 73.50 Matches: 40
Percent Similarity: 38.71% Conservative: 32
Best Local Similarity: 21.51% Mismatches: 65
Query Match: 7.68% Indels: 49
DB: 9 Gaps: 12
US-09-995-493-52 (1-179) x US-10-114-170-174 (1-7304)
QY 16 ThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGln-----Leu 31
Db 6358 ACGATCAACAGGAGAGCTAC-----TCAGCTTATGCACGAGTGGCTCTATCTG 6408
QY 32 ThrAspGlnTrp-----AlaIleHisTyrSerTyrLysTyrTrpGluTrpSer 47
Db 6409 ACCGATAACTGGATTGCGGTCGGCGGATCCGCTATCAGTAT--TACAGCGCAGTATGCG 6465
QY 48 ArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLys 67
Db 6466 -----GGCGGTCCTTTTAAATGTCAATACTGAC 6498
QY 68 LysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAsp 87
Db 6499 AGCCGCGATGAACAAATGGAGCCCAAA-----CTGGGGTTAGTCTACAAACTGACGCCA 6552
QY 88 AlaLeuThrLeuArgAlaGlyLeuAla-----TyrAspLysAlaAlaSerLysThrHis 105
Db 6553 TCGGTATCCTTATTGCTAATTATTCGCAACATTTATCCGCAATCGTCAATGGCCAGC 6612
QY 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125

Db 6613 TACATCGGAGATCTTCACCGGAATCATCTAATGCTTACGAAGTCGGGGCAAAATTCGAG 6672
QY 126 PheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLysHis 145
Db 6673 CTGTTCCGATGGTATAACCGCAGATATTGCGCTGTTTATATCCAT-----AAACGTAAC 6726
QY 146 PheValGluThrGlnAsnIle----- 152
Db 6727 GTCTTGTATACCGAAAGATTGTTGATGAACCATGCCAAAACGGCAGCCGGTTCGT 6786
QY 153 ---LysGlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeu--- 170
Db 6787 TCAAGAGGG-----GTAGAAGTCGACCTTTGCGGGAGCATTAATACTGAACATTAAT 6837
QY 171 -----TyrGly 172
Db 6838 ATCATTGCCAGCTACGGC 6855
RESULT 27
US-09-850-351A-33/C
Sequence 33, Application US/09850351A
Patent No. US20020100080A1
GENERAL INFORMATION:
APPLICANT: Feitelson, Jerald S.
Schnef, H. Ernest
Narva, Kenneth E.
Stockhoff, Brian A.
Schmeits, James
Loewer, David
Dulium, Charles Joseph
Mueller-Cohn, Judy
Stamp, Lisa
Morrill, George
TITLE OF INVENTION: No. US20020100080A1e1 Pesticidal Toxins and Nucleotide
Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/850,351A
FILING DATE: 07-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/073,898
FILING DATE: 06-May-1998
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-Oct-1997
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-Oct-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708CD1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 17718
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-850-351A-33

Alignment Scores:
Pred. No.: 3,26 Length: 1022
Score: 73.00 Matches: 35
Percent Similarity: 38.40% Conservatve: 13
Best Local Similarity: 28.00% Mismatches: 49
Query Match: 7.63% Indels: 29
DB: 10 Gaps: 7

US-09-995-493-52 (1-179) x US-09-850-351A-33 (1-1022)

```
Qy 29 HisGlnLeuThrAspGlnTrp---AlaIleHsTyrSerTyrLysTyrThrGluTrp--- 46
|||||
Db 878 CACGAGCTT-----CAATGGAAGCTCTTCTGATTTCGTATAAGACCAATTCGTGGATG 825
Qy 47 SerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThr 66
|||
Db 824 AATGAGAC-TCTACATATTGGATAAATTTTCATTGGT----- 787
Qy 67 LysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsn 86
|||||
Db 786 -----GATAATATCACCTTTTCCATA---CTAACATTTCACACTTTGGR 748
Qy 87 AspAlaLeuThr-----LeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThr 104
|||||
Db 747 AAGCAGCTACCAATGGGTTGAAGCTTTCCCTTTCGATTTCGATAAATCTAAATCCCTTGGC 688
Qy 105 HisLeuSerAlaSerIleProAspThrAspArgMetTyrSerIleGlyAlaThrTyr 124
|||||
Db 687 GCCTTTTCATATCATAGTAGGATCCCAACTGTGTGGCTGTCTAATGGATTCGAACA 628
Qy 125 LysPheThrProAsnLeuSerValaspValGlyPheAlaHsLeuArgGlyLysLys 144
|||
Db 627 AATTTTGTA-----TATCCCTTACTTGTCTAGCGAATCATCC 592
Qy 145 HsPheValGluThr 149
|||
Db 591 CATTGGACAGCAACT 577
```

RESULT 28

US-09-764-891-10186/c
Sequence 1, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10186
LENGTH: 15261
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-10186

Alignment Scores:
Pred. No.: 170 Length: 15261
Score: 73.00 Matches: 49
Percent Similarity: 33.98% Conservatve: 21
Best Local Similarity: 23.79% Mismatches: 81
Query Match: 7.63% Indels: 55
DB: 9 Gaps: 10

US-09-995-493-52 (1-179) x US-09-764-891-10186 (1-15261)

```
Qy 12 LysGlySerLeuThrLeuLysLeuProAlaTyrTrpGlu-LeuSerGlyPheHs----- 29
|||
Db 10993 AAGCAATCCTCCTGGCTTGGCTCCCAAGTGTGGGAATACAGCGCTGAGCCACTGTGC 10934
Qy 30 -----GlnLeuThrAspGlnTrpAlaIleHsTyrSe 40
|||||
Db 10933 CCAGCCCAATGCTAATAACATGAAGAAACCCACTGAGTGACCATTCCTAGAGACACTATCA 10874
Qy 40 rTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySe 60
|||
Db 10873 TGTCAACTTGAGACCTTGGGAGACACAGAAA-----AACAAATATAAGAT----- 10828
Qy 60 rGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleG1 80
|||
Db 10827 -----AGAATTACAGGCAAGAGACGTAATCTCTGGGGCAAGGACCTTCCTCAA 10778
Qy 80 yThrThrTyrSerLeuAsn----- 86
|||
Db 10777 CACTTCTTAACCAAAATAAGTGGCTTATTAGATGCATTTCTTGGCTGCAGAGGAAAAAT 10718
Qy 87 -----AspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAsp----- 98
|||
Db 10717 GATAAGCAACTCACCCCTCCCATCTCAGATTGAGGGGTGTCTTGGTAATGACAAATTGTCT 10658
Qy 99 ----LysAlaAlaSerLysThrHsLeuSerAlaSerIlePro-----AspThrAs 114
|||
Db 10657 TCCCAAGCTTGGATTCCCAAGTCCACCTTGTAGTCCCTATGCCGTAAAAAGCTCTCGCTTGA 10598
Qy 114 pArgMet---TrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValas 133
|||
Db 10597 CAGCTGCATCTGGTTTCTCGAAACATAATATGGGGAAGGAAGTCCCTTCAATA----- 10546
Qy 133 pValGlyPheAlaHsLeuArgGlyLysLysLysPheValGluThrGlnAsnIleLy 153
|||
Db 10545 -TTATAGTTCTCTCAACAGCAAGAAAGAAAAAATGCATGAATCCAAATTCAGGCTCTTAC 10487
Qy 153 sGlyLeuLeuLeuVal-----GluAlaAspTyrThrLysAlaThrAlaAsnLeuTy 171
|||
Db 10486 TCATTTCAAATCATCTCTGAAAGGACACACAGACACTTAAAGTAGCAGCT----- 10435
Qy 171 rGlyLeuAsnLeuAsn 176
|||
Db 10434 -GGGAGAATCTCAAC 10420
```

RESULT 29

US-09-913-514-1
Sequence 1, Application US/09913514
Publication No. US20030082210A1
GENERAL INFORMATION:
APPLICANT: GOMI, Yasuyuki
APPLICANT: SUNAMACHI, Hiroki
APPLICANT: TAKAHASHI, Michiaki
APPLICANT: YAMANISHI, Koichi
TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vac
FILE REFERENCE: 0216-0454P
CURRENT APPLICATION NUMBER: US/09/913,514
CURRENT FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: PCT/Jp01/00678
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: JP 2000-62734
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 124884
TYPE: DNA
ORGANISM: Varicella vlrus
FEATURE:
NAME/KEY: mls_cfeature
LOCATION: (1)..(124884)
OTHER INFORMATION: Dumas Strain
US-09-913-514-1


```
Alignment Scores:
Pred. No.: 3.7e+03 Length: 124884
Score: 73.00 Matches: 41
Percent Similarity: 38.51% Conservative: 16
Best Local Similarity: 27.70% Mismatches: 42
Query Match: 7.63% Indels: 50
DB: 9 Gaps: 9

US-09-995-493-52 (1-179) x US-09-913-514-1 (1-124884)
QY 10 IleGlyLysGlySerLeu-----ThrLeuLysLeu-----ProAlaTyrTrp 23
DB 39637 TTAGGAACCCAGATCTTCTATAGACGCTTGTGCGAAAAAATTAAACCTTTTACTTGG 39696
QY 24 GluLeuSerGlyPheHis-----GlnLeu 31
DB 39697 ATACTGATGGGTTCACCTTCGATATAGACTGTGTATATAATCAGTTGGGTTAAACTTG 39756
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu----- 45
DB 39757 GAACCATGCTATGACGTACAGGTGGATTATCTTTACGACCGGCACCTTCCCGATGAAGGTA 39816
QY 46 ---TrpSer-----ArgPheLysGluLeuArg 53
DB 39817 TTGTGGAAGAACTACACTATCGGCATTCAGCGTGCCCAATGAGGTC-AATGAGCTACGC 39875
QY 54 GlyLysTyrGlnAspGly-----SerGlyTyrGluAlaPheThr----- 66
DB 39876 ATTGAATAGCATCCGCTATAAATCCGGGTTTCTGCCTTTTCCACCTTTGTTAGGCAT 39935
QY 67 ---LysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSer--- 84
DB 39936 CGCCACCGCGAATGGGTAAACCAACGACGACGAGCCATTGCGAGATATACGCCGCGC 39995
QY 85 ---LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLys-----Ala 100
DB 39996 CTTATAACCAACACATTTGACACGACATACGGGGTTTCATTGGGACAAAGCTTATTATTCT 40055
QY 101 AlaSerLysThrHisLeuSerAla 108
DB 40056 TTTGAAAAACACCCTTAACCTCT 40079

RESULT 30
US-09-913-514-2
; Sequence 2, Application US/09913514
; Publication No. US20030082210A1
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAMACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccin
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JP01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 125157
; TYPE: DNA
; ORGANISM: Varicella vlrus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(125157)
; OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

Alignment Scores:
Pred. No.: 3.71e+03 Length: 125157
Score: 5.14 Matches: 1254
Percent Similarity: 72.50 Conservative: 45
Best Local Similarity: 33.82% Mismatches: 25
Query Match: 21.74% Indels: 66
DB: 10 Gaps: 12

US-09-995-493-52 (1-179) x US-09-925-297-58 (1-125157)
QY 18 LysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIle 37
```

```
Score: 73.00 Matches: 41
Percent Similarity: 38.51% Conservative: 16
Best Local Similarity: 27.70% Mismatches: 42
Query Match: 7.63% Indels: 50
DB: 9 Gaps: 9

US-09-995-493-52 (1-179) x US-09-913-514-2 (1-125157)
QY 10 IleGlyLysGlySerLeu-----ThrLeuLysLeu-----ProAlaTyrTrp 23
DB 39588 TTAGGAACCCAGATCTTCTATAGACGCTTGTGCGAAAAAATTAAACCTTTTACTTGG 39647
QY 24 GluLeuSerGlyPheHis-----GlnLeu 31
DB 39648 ATACTGATGGGTTCACCTTCGATATAGACTGTGTATAAATCAGTTGGGTTAAACTTG 39707
QY 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu----- 45
DB 39708 GAACCATGCTATGACGTACAGGTGGATTATCTTTACGACCGGCACCTTCCCGATGAAGGTA 39767
QY 46 ---TrpSer-----ArgPheLysGluLeuArg 53
DB 39768 TTGTGGAAGAACTACACTATCGGCATTCAGCGTGCCCAATGAGGTC-AATGAGCTACGC 39826
QY 54 GlyLysTyrGlnAspGly-----SerGlyTyrGluAlaPheThr----- 66
DB 39827 ATTGAATAGCATCCGCTATAAATCCGGGTTTCTGCCTTTTCCACCTTTGTTAGGCAT 39886
QY 67 ---LysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSer--- 84
DB 39887 CGCCACCGCGAATGGGTAAACCAACGACGACGAGCCATTGCGAGATATACGCCGCGC 39946
QY 85 ---LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLys-----Ala 100
DB 39947 CTTATAACCAACACATTTGACACGACATACGGGGTTTCATTGGGACAAAGCTTATTATTCT 40006
QY 101 AlaSerLysThrHisLeuSerAla 108
DB 40007 TTTGAAAAACACCCTTAACCTCT 40030

RESULT 31
US-09-925-297-58
; Sequence 58, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-297-58

Alignment Scores:
Pred. No.: 5.14 Length: 1254
Score: 72.50 Matches: 45
Percent Similarity: 33.82% Conservative: 25
Best Local Similarity: 21.74% Mismatches: 66
Query Match: 7.58% Indels: 71
DB: 10 Gaps: 12

US-09-995-493-52 (1-179) x US-09-925-297-58 (1-1254)
QY 18 LysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIle 37
```

```
Db 517 AAGCGTCAGCCATCTCGATCGACACGGCATCCAT-----TCCCGGAGTGGGTCAACC 570
Qy 38 HlStyrsSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
Db 571 CAGGCCAGT-----GGGCTCTG-----589
Qy 58 AspGlySerGlyTyrGluAlaPheThrLysLys-----GluGluTyrLysAspAsnSer 75
Db 589 -----TTTGCAAGAAGATCACTCAAGACTACGGCAGGATGCA 627
Qy 76 ArgPheAla-----IleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArgAla 93
Db 628 GCTTTTCCGCCATTCCTCGACACCTTCCTGGAGATCGTCACCAACCCCTGAT 687
Qy 94 GlyLeuAlaTyrAspLysAlaAla-----SerLysThrHisLeuSer 107
Db 688 GGCTTTTCCGCCATTCCTCGACACCTTCCTGGAGATCGTCACCAACCCCTGAT 747
Qy 108 AlaSerIle-----ProAspThrAspArgMetTrpTyrSer-----119
Db 748 GGCTCCCTCTGATTGGCTGGACCCCAACAGGAACCTGGGCGCTTGGGTTGTC 807
Qy 120 -----IleGlyAlaThrTyr-----LysPheThrProAsnLeuSer 131
Db 808 GGAGCCAGCAGTAACCCCTGCTCGAGACTTACCACGGCAAGTTTGCC---AATTCCGAA 864
Qy 132 ValAsp-----ValGlyPheAlaHisLeuArgGlyLysLysHisPheVal 147
Db 865 GTGAGGTCAAGTCCATTGTAGACTTTGTGAAGGACCATGGACATCAAGGCCTTCATC 924
Qy 148 GluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrLys-----165
Db 925 TCCATCCACAGCTACTCCAGCTCCTCATGTATCCCTATGGCTACAAACAGAACAGTC 984
Qy 166 -----AlaThrAlaAsnLeu 170
Db 985 CTGACCAGGATGAGTGGATCAGCTTTTCCAGGCTCCTGTGACAGCGCTGGCTCTCTC 1044
Qy 171 TyrGlyLeuAsnLeuAsnTyr 177
Db 1045 TACGGGACCAAGTTCAACTAT 1065
```

RESULT 32

```
US-09-969-347-209
; Sequence 209, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 209
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-209
```

Alignment Scores:

Pred. No.:	5.92	Length:	1380
Score:	72.50	Matches:	45
Percent Similarity:	33.82%	Conservative:	25
Best Local Similarity:	21.74%	Mismatches:	66
Query Match:	7.58%	Indels:	71
DB:	10	Gaps:	12

```
US-09-995-493-52 (1-179) x US-09-969-347-209 (1-1380)
Qy 18 LysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIle 37
Db 509 AAGCGTCAGCCATCTCGATCGACACGGCATCCAT-----TCCCGGAGTGGGTCAACC 562
Qy 38 HlStyrsSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
Db 563 CAGGCCAGT-----GGGCTCTG-----580
Qy 58 AspGlySerGlyTyrGluAlaPheThrLysLys-----GluGluTyrLysAspAsnSer 75
Db 581 -----TTTGCAAGAAGATCACTCAAGACTATGGCAGGATGCA 619
Qy 76 ArgPheAla-----IleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArgAla 93
Db 620 GCTTTTCCGCCATTCCTCGACACCTTCCTGGAGATCGTCACCAACCCCTGAT 679
Qy 94 GlyLeuAlaTyrAspLysAlaAla-----SerLysThrHisLeuSer 107
Db 680 GGCTTTTCCGCCATTCCTCGACACCTTCCTGGAGATCGTCACCAACCCCTGAT 739
Qy 108 AlaSerIle-----ProAspThrAspArgMetTrpTyrSer-----119
Db 740 GGCTCCCTCTGATTGGCTGGACCCCAACAGGAACCTGGGCGCTTGGGTTGTC 799
Qy 120 -----IleGlyAlaThrTyr-----LysPheThrProAsnLeuSer 131
Db 800 GGAGCCAGCAGTAACCCCTGCTCGAGACTTACCACGGCAAGTTTGCC---AATTCCGAA 856
Qy 132 ValAsp-----ValGlyPheAlaHisLeuArgGlyLysLysHisPheVal 147
Db 857 GTGAGGTCAAGTCCATTGTAGACTTTGTGAAGGACCATGGACATCAAGGCCTTCATC 916
Qy 148 GluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThrLys-----165
Db 917 TCCATCCACAGCTACTCCAGCTCCTCATGTATCCCTATGGCTACAAACAGAACAGTC 976
Qy 166 -----AlaThrAlaAsnLeu 170
Db 977 CTGACCAGGATGAGTGGATCAGCTTTTCCAGGCTCCTGTGACAGCGCTGGCTCTCTC 1036
Qy 171 TyrGlyLeuAsnLeuAsnTyr 177
Db 1037 TACGGGACCAAGTTCAACTAT 1057
```

RESULT 33

```
US-09-974-300-4357
; Sequence 4357, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4357
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-4357
```

Alignment Scores:

Pred. No.:	5.91	Length:	1239
------------	------	---------	------

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 212
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-212

Alignment Scores:
Pred. No.: 3.69 Length: 807
Score: 71.50 Matches: 37
Percent Similarity: 39.13% Conservative: 26
Best Local Similarity: 22.98% Mismatches: 55
Query Match: 7.47% Indels: 43
DB: 9 Gaps: 7

US-09-995-493-52 (1-179) x US-09-938-842A-212 (1-807)
Qy 27 GlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrp 46
Db 22 GGTTCCTCAGTGTGGAAGGCAAG---TCCAGCCATCCGATGGAGGACTATGTAGTG--- 75
Qy 47 SerArgPheLysGluLeuArgGly----- 54
Db 76 TCTGAATTCAGAACTTGAAGGTCATGAATGGGTTTGTGCTATCTTTGATGGTCAC 135
Qy 55 -----LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLys 68
Db 136 TTGGGCGATGATGGCTAAATCTTCAGACTAATCTCTTGACAAACATCTTTAAAGAG 195
Qy 69 GluGluTyrLysAspAsnSerArgPheAlaIleGlyThrTyrThrTyrSerLeuAsnAspAla 88
Db 196 AAGGATTTTGGAGTACACAGAGATGCTATAGGAATGCCTACAGATCAACAGATGCC 255
Qy 89 LeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSer----- 102
Db 256 GTGATATTGACGAGTCCCTTAAGCTTGGTAAGCGCGATCAACAGCTGTAAACGGGAAT 315
Qy 103 -----LysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyr 118
Db 316 CTAATGATGCTAAAGAGCTAGTTGTTGCTAATGTTGACAGCTCCAGAGCATGATGCT 375
Qy 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
Db 376 AAGATGGTGTGGCCATCAG-----CTATCTGTCTGAT-----CAT 411
Qy 139 LeuArgGlyLysLysHisPheValGluThrGln-----AsnIleLys 153
Db 412 GAACCAAGCAAGGAGAAAGAAAGAAATAGAGCGGAGTGGCTTTGTGTCAAATATTCCA 471
Qy 154 Gly 154
Db 472 GGG 474

RESULT 36
US-09-974-300-1094
; Sequence 1094, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berkas, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526

; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1094
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1094

Alignment Scores:
Pred. No.: 3.05 Length: 636
Score: 71.00 Matches: 27
Percent Similarity: 43.40% Conservative: 19
Best Local Similarity: 25.47% Mismatches: 40
Query Match: 7.42% Indels: 20
DB: 10 Gaps: 5

US-09-995-493-52 (1-179) x US-09-974-300-1094 (1-636)

Qy 54 GlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAsp 73
Db 181 GGAATTCGAGTCGGGCTGGGCTTTGAACAGCTTTCTGAACACGAAAGCTAAA--- 237
Qy 74 AsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThrLeuArgAla 93
Db 238 -----GTCCGCCGCTTGAATGAGCTGAGCGAGCAATTT--- 270
Qy 94 GlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeu---SerAlaSerIleProAsp 112
Db 271 ---ATCCGGGAGACAATAGCTGTTGTCACGCCGCTTTGGGAATTTCTCTTTCCCGCG 327
Qy 113 ThrAspArgMetTrpTyrSer-----IleGlyAlaThrTyrLysPheThrPro 128
Db 328 GTGATGAAGAGCTACTTCGATGCGTGTAGCGGAAAGAGCTTTAAATATACCGAA 387
Qy 129 -----AsnLeuSerValAspValGlyPheAlaHisLeuArgGlyLysLys 144
Db 388 CAGGACCGATCGGGCTGTAAACGATAAAGAGCGCTGCACATCCAGCGAGCGCGC 447
Qy 145 HisPheValGluThrGln 150
Db 448 TACTATTCAAGGCGCAG 465

RESULT 37
US-10-114-170-86
; Sequence 86, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod

TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Plunkett Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION NUMBER: 09/453,702

RESULT 38
US-09-822-849A-513
; Sequence 513, Application US/09822849A

```
Db 511 ACCCATGTGCC-----AAGGTTGGATC----- 534
Qy 164 ThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyr 177
Db 535 ---AAGGCAATGCCATGACATCGCGATGGATTATGATTAT 573

RESULT 39
US-09-738-626-216/c
; Sequence 216, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 216
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-216

Alignment Scores:
Pred. No.: 12.1 Length: 1464
Score: 70.50 Matches: 31
Percent Similarity: 38.10% Conservative: 25
Best Local Similarity: 21.09% Mismatches: 62
Query Match: 7.37% Indels: 29
DB: 9 Gaps: 6

US-09-995-493-52 (1-179) x US-09-738-626-216 (1-1464)
Qy 11 GlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGln 30
Db 951 GGAAAGGTCAGCTT---GTTCAAGTCCAGGATGGGAAGATATCTGGAACCTTTCATCAG 895
Qy 31 LeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGlu---TrpSerArgPhe 49
Db 894 CTG-----TGGAGCATTAACCTTCAGTGGATCTCTGGCCAGCTTTGGTGCCTTC 844
Qy 50 LysGluLeu-----ArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPhe 65
Db 843 TTCAGCTGACCAAGTGTGCCAGGGAACACCGAGATGGTGGAGCGGCCAAGCTTTGG 784
Qy 66 ThrLysLysGluGluTyrLysAspAsnSerArgPhe----- 77
Db 783 CTCAGGCGTGATGATCGAACCACCGAAGAAATAGGTGTAGTTGTCTGTATCGCGTC 724
Qy 78 -----AlaIleGlyThrThrTyrSerLeuAsnAspAla 88
Db 723 GGAGTGGACCAACGGTTTCTTCCAAAGTTGTTGGTGGCGTGGTCCGCAATGAAGTA 664
Qy 89 LeuThrLeuArgAlaGlyLeuAla-----TyrAspLysAlaAlaSerLysThrHis 105
Db 663 GCGCGTTTCGCTCTGTGTCATCGCGATCGGTGACGACGATGATGCGGTGTCAGGCCCAT 604
```

```
Qy 106 LeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLys 125
Db 603 GCCACCAACGTTGCCCAAGACAGGTCCGCTGTGGGTCTTCGGCGGTGCTTCTGGCTC 544
Qy 126 PheThrProAsnLeuSerVal 132
Db 543 GAGGTGCAGGATCGTCCGTC 523

RESULT 40
US-09-765-205-11
; Sequence 11, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
; APPLICANT: CAO, LI
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458.004/200130.449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 1585
; TYPE: DNA
; ORGANISM: human
US-09-765-205-11

Alignment Scores:
Pred. No.: 13.6 Length: 1585
Score: 70.50 Matches: 51
Percent Similarity: 34.11% Conservative: 22
Best Local Similarity: 23.83% Mismatches: 66
Query Match: 7.37% Indels: 75
DB: 10 Gaps: 11

US-09-995-493-52 (1-179) x US-09-765-205-11 (1-1585)
Qy 7 GlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrpGluLeuSer 26
Db 236 GGACCCCAAGTGTCAATAGGAACT-----CCATGCCCACTTACGAAGAGGCCAAG 286
Qy 27 GlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrp 46
Db 287 CAATAT-----CTGTCTTTATGAAGAGCTCTATGCCAATGGC 322
Qy 47 SerArgPheLysGluLeuArgGlyLysTyr----- 56
Db 323 AGCCGCACAGACAGCGCAGGTGGGCTACATCTCAGCAGTAGTGGAGATGGGGCCAA 382
Qy 57 -----GlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsn 74
Db 383 CACCGACACTCAGGCTCTTCAGGAAAGCTCTCGAAGAAAGCGGCAAGATTATGGCTATGAC 442
Qy 75 SerArgPheAlaIle---GlyThrThrTyrSerLeuAsn-----AspAlaLeu 89
Db 443 AGCAGGTTTCAGCATTTTGGGAAGGACTTCTGCTCAACTACCTTCTCAACATCAGTG 502
Qy 90 ThrLeuArgAlaGlyLeuAlaTyrAspLysAlaSerLysThrHisLeuSerAlaSer 109
Db 503 AAGTTATCCAGGCTCCACCGGCCACCCCTGGTGGCAGAGAGCATGTCCTCACAGCTGCC 562
Qy 110 -----IleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyr---LysPhe 126
Db 563 CACTGCATACAGGAT-----GGAAAAACCTATGTGAAGGA 598
Qy 127 ThrProAsnLeuSerVal----- 132
Db 599 ACCCAGAAAGTTCGAGTGGGCTTCCTAAAGCCCAAGTTTAAAGATGGTGTGTCAGGGGCC 658
Qy 133 -----AspValGlyPheAlaHisLeuArgGlyLysLys 143
Db 659 AACGACTCCACTTCAGCCATGCCCGACGACAGATGAATTTTCAGTGGATCCGGGTGAACGC 718
```

Qy 144 LysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThr 163
Db 719 ACCCATGTGCCC-----AAGGTTGGATC-----742

Qy 164 ThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAsnTyr 177
Db 743 ---AAGGGCAATCCCAATGACATCGGCATGGATTATGATTAT 781

Search completed: May 19, 2003, 20:36:41
Job time : 174 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 19, 2003, 19:03:23 ; Search time 1078 Seconds
(without alignments)
2689.232 Million cell updates/sec

Title: US-09-995-493-52
Perfect score: 957
Sequence: 1 OHNGVLGPYIGKSLTLKLP.....ADYTTKATANLYGLNLNRYF 179

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US0995493/runat_12052003_091309_23015/app_query.fasta_1.327
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0995493.ecgn_1.1456/runat_12052003_091309_23015 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGIO -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOB=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	304	31.8	997	17	CNS010FK
c 2	120.5	12.6	796	17	BH385201
c 3	100.5	10.5	746	17	BH385201 AG-ND-160
c 4	96.5	10.1	659	13	BH383328 AG-ND-160
c 5	96.5	10.1	727	13	BJ429628 BJ429628
c 6	96.5	10.1	747	14	BJ429792 BJ429792
c 7	86	9.0	621	17	C91402 C91402 Dict
c 8	85	8.9	941	17	AQ990154 RfC00875
c 9	84	8.8	941	17	BQ706806 AGENCOURT
c 10	83.5	8.7	431	13	B1547625 G03191761
c 11	83.5	8.7	756	12	B1139361 F130P15Y
c 12	83	8.7	699	10	BG064935 H3025E07-
c 13	82.5	8.6	600	17	BE005948 RCO-BN012
c 14	82.5	8.6	649	10	BH878604 hs83h03 b
c 15	82.5	8.6	683	13	AW55347 L0227C06-
c 16	82.5	8.6	744	14	BJ041818 BJ041818
c 17	82	8.6	665	17	BM930833 UT-E-EJ1-
c 18	82	8.6	725	13	AZ573577 319PVA06
c 19	81.5	8.5	386	10	BM396839 5009-0-25
c 20	81.5	8.5	417	10	AW788126 C00212-F
c 21	81.5	8.5	563	13	AW759240 G138f12.y
c 22	81.5	8.5	1496	11	BM178271 saj71a04.
c 23	81	8.5	659	12	AK010501 Mus muscu
c 24	81	8.5	873	9	BG740114 G02631193
c 25	80.5	8.4	484	14	AL531424 AL531424
c 26	80.5	8.4	484	14	W64749 md88c04.r1
c 27	80	8.4	516	14	W64824 md88g04.r1
c 28	80	8.4	571	9	BQ36341 md07g02.y
c 29	80	8.4	598	13	AL701828 DKFZP686H
c 30	80	8.4	599	13	BI819893 G03035435
c 31	80	8.4	611	9	BM090644 lg16g01.y
c 32	80	8.4	615	12	AL531052 AL531052
c 33	80	8.4	617	14	BG700910 G02681855
c 34	80	8.4	623	14	BQ639058 hd30h04.y
c 35	80	8.4	631	14	BQ639009 hd30c08.y
c 36	80	8.4	647	14	BQ639930 he22a07.y
c 37	80	8.4	647	14	BM707016 UT-E-CQ0-
c 38	80	8.4	650	13	BQ636248 hd06d06.y
c 39	80	8.4	662	14	BI669859 G03293455
c 40	80	8.4	665	10	BQ635968 hd02f11.y
c 41	80	8.4	707	12	BE006001 RCO-BN012
c 42	80	8.4	715	12	BG829263 G02753162
c 43	80	8.4	739	12	BG705579 G02668762
c 44	80	8.4	780	13	BG708348 G02672287
c 45	80	8.4	791	13	BI549564 G03192261
					BM399133 5009-0-53

ALIGNMENTS

RESULT 1
CNS010FK/c
LOCUS
DEFINITION
CNS010FK
Anopheles gambiae GSS SP6 end of clone 26E01 of Notredame library
from strain PEST of Anopheles gambiae (African malaria mosquito).
genomic survey sequence.
AL153456
AL153456.1 GI:7014375
GSS.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE
1 (bases 1 to 997)

[illegible]

BASE COUNT 231 a 118 c 90 g 220 t
ORIGIN

Alignment Scores:

Pred. No.: 0.0781 Length: 659
Score: 96.50 Matches: 39
Percent Similarity: 42.47% Conservative: 23
Best Local Similarity: 26.71% Mismatches: 53
Query Match: 10.08% Indels: 31
DB: 13 Gaps: 7

US-09-995-493-52 (1-179) x BJ429628 (1-659)

QY 38 HistyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
|||||
DB 418 CATATTTCGTAT-----AGTAATATCGT 395

QY 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
|||||
DB 394 AATGGAATTCATATTTGGCATTTGGTTTAAAGAAGCTTTTAGAATGTATCCAGCTGGT 335

QY 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
|||||
DB 334 ATTTAGTTTACCTCATATGACTTCTGAGGATGTCGAGATTCAAGGT-----287

QY 98 AspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrp 117
|||||
DB 286 CATAAATTCGAAGGCTACTCAATTCANAAATAT-----248

QY 118 TyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAla 137
|||||
DB 247 TACTCTACTCATGCTAGTGAATCTTTTGGCCAAATCCAAATATTTATACACCAAGAAGA 188

QY 138 HisLeuArg-----GlyLysLysLysHisPhe--ValGluThrGlnAsnIleLys 153
|||||
DB 187 CATATCAAAATGATGTTCTAAAGTGTTCATTTCCCTGTTGGTAGAAATTTGATG 128

QY 154 GlyLeuLeuValGluAlaAspTyrThrLysAlaThrLysAlaAsnLeuTyrGlyLeu 173
|||||
DB 127 GGTATGAGTTAAGTGAGGCTGAAGTTCATACT---GCTATGGCTGAATTTATTTGGT---74

QY 174 AsnLeuAsnTyrArgPhe 179
|||||
DB 73 -----AATTTCAAATTT 62

RESULT 5
BJ429792/c

LOCUS BJ429792 727 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ429792 Dictyostellium discoideum cDNA library, VF Dictyostellium
discoideum cDNA clone ddv4k24 3', mRNA sequence.

ACCESSION BJ429792.1 GI:19404514
VERSION
KEYWORDS
SOURCE Dictyostellium discoideum.

REFERENCE Eukaryota; Mycetozoa; Dictyostelid; Dictyostellium.
1 (bases 1 to 727)

AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostellium discoideum at the vegetative
stage

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source

1. .727
/organism="Dictyostellium discoideum"
/strain="AX4"

/db_xref="taxon:44689"
/clone="ddv4k24"
/clone_lib="Dictyostellium discoideum cDNA library, VF"
/sex="mat A"
/dev_stage="Growth phase"
BASE COUNT 245 a 124 c 98 g 256 t 4 others
ORIGIN

Alignment Scores:

Pred. No.: 0.0909 Length: 727
Score: 96.50 Matches: 39
Percent Similarity: 42.47% Conservative: 23
Best Local Similarity: 26.71% Mismatches: 53
Query Match: 10.08% Indels: 31
DB: 13 Gaps: 7

US-09-995-493-52 (1-179) x BJ429792 (1-727)

QY 38 HistyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
|||||
DB 460 CATATTTCGTAT-----AGTAATATCGT 437

QY 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
|||||
DB 436 AATGGAATTCATATTTGGCATTTGGTTTAAAGAAGCTTTTAGAATGTATCCAGCTGGT 377

QY 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
|||||
DB 376 ATTTAGTTTACCTCATATGACTTCTGAGGATGTCGAGATTCAAGGT-----329

QY 98 AspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrp 117
|||||
DB 328 CATAAATTCGAAGGCTACTCAATTCANAAATAT-----290

QY 118 TyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAla 137
|||||
DB 289 TACTCTACTCATGCTAGTGAATCTTTTGGCCAAATCCAAATATTTATACACCAAGAAGA 230

QY 138 HisLeuArg-----GlyLysLysLysHisPhe--ValGluThrGlnAsnIleLys 153
|||||
DB 229 CATATTCAAATGATGTTTCTAAAGTGTTCATTTCCCTGTTGGTAGAAATTTGATG 170

QY 154 GlyLeuLeuValGluAlaAspTyrThrLysAlaThrLysAlaAsnLeuTyrGlyLeu 173
|||||
DB 169 GGTATGAGTTTAAAGTGAGGCTGAAGTTCATACT---GCTATGGCTGAATTTATTTGGT---116

QY 174 AsnLeuAsnTyrArgPhe 179
|||||
DB 115 -----AATTTCAAATTT 104

RESULT 6
C91402

LOCUS C91402 747 bp mRNA linear EST 20-APR-1998
DEFINITION C91402 Dictyostellium discoideum SS (H. Urushihara) Dictyostellium
discoideum cDNA clone SSK171, mRNA sequence.

ACCESSION C91402.1 GI:3060768
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Dictyostellium discoideum.

REFERENCE Dictyostellium discoideum; Dictyostelid; Dictyostellium.

AUTHORS

1 (bases 1 to 747)
Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
Yoshino, R., Mitra, B. N., Pi, M., Sato, T., Takenoto, K., Yasukawa, H.,
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

TITLE
JOURNAL

COMMENT Unpublished (1998)
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp.
 Location/Qualifiers
 1. 747
 /organism="Dictyostellium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SSK171"
 /dev_stage="slug"
 298 a 81 c 116 g 252 t
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 0.0948 Length: 747
 Score: 96.50 Matches: 39
 Percent Similarity: 42.4% Conservative: 23
 Best Local Similarity: 26.71% Mismatches: 53
 Query Match: 10.08% Indels: 31
 DB: 14 Gaps: 7

US-09-995-493-52 (1-179) x C91402 (1-747)
 QY 38 HisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGln 57
 Db 202 CATTATTCGTAT-----AGTAAATATCGT 225
 QY 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
 Db 226 ATGGAAATTCATATTTGGCATGTGTTTAAGAAGATTTTATAGATGATCCAGCTGGT 285
 QY 78 AlaIleGlyThrTyrThrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLcuAlaTyr 97
 Db 286 ATTTAGGTTTACCTCATATGACTTCTGAGGATTCGAGATTCAAGT----- 333
 QY 98 AspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrp 117
 Db 334 CATAAATTTGCAAGGGTACTCAAAATTCATTCAAATATT----- 372
 QY 118 TyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAla 137
 Db 373 TACTCTACTCATCGTAGTAACTTTTGGCCAAATCCAAATAATTTATACCAGAAAGA 432
 QY 138 HisLeuArg-----GlyTysLysLysHisPhe---ValGluThrGlnAsnIleLys 153
 Db 433 CATATTCAAAATGATCTTTCTAAAGTGTTCATTTCGCTGTTGCTACTAGAAATTCATG 492
 QY 154 GlyLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeu 173
 Db 493 GGTATGAGTTTAAGTCAGGCTGAAGTTCATCTACT---GCTATGCTCAATATTATTTGGT--- 546
 QY 174 AsnLeuAsnTyrArgPhe 179
 Db 547 -----AATTTCAATTT 558

RESULT 7
 A0990154/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Photobacterium.
 1 (bases 1 to 621)
 ffrench-Constant.R.H., Waterfield.N., Burland.V., Perna.N.T.,
 Daborn.P.J., Bowen.D. and Blattner.F.R.
 A genomic sample sequence of the entomopathogenic bacterium
 Photobacterium luminescens W14: potential implications for virulence
 Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
 20378633

Contact: ffrench-Constant RH
 Department of Biology and Biochemistry
 University of Bath
 South Building, Bath BA2 7AY, UK
 Tel: (44) 1225 826621
 Fax: (44) 1225 826779
 Email: bssr@bath.ac.uk
 This is one of 2,122 random reads from the M13 library. For
 annotation of identified clones (BLASTX, BLASTN and mapping to E.
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
 Acids Res.
 Seq primer: M13 Forward
 Class: shotgun.
 Location/Qualifiers
 1. 621
 /organism="Photobacterium luminescens"
 /strain="W14"
 /db_xref="taxon:29488"
 /clone="PLG00876"
 /clone_lib="Photobacterium luminescens strain W14 M13
 library"
 /dev_stage="primary phase variant"
 /note="genomic DNA from strain W14 was size selected (1-2
 kb) and then cloned into M13 Janus."
 BASE COUNT 183 a 111 c 109 g 216 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.61 Length: 621
 Score: 86.00 Matches: 38
 Percent Similarity: 37.80% Conservative: 24
 Best Local Similarity: 23.17% Mismatches: 59
 Query Match: 8.99% Indels: 43
 DB: 17 Gaps: 7

US-09-995-493-52 (1-179) x A0990154 (1-621)
 QY 29 HisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArg 48
 Db 539 TATGACTTCATGATCATCAATGGGGCTATATCATCATTTACNTTTCCT----- 492
 QY 49 PheLys-GluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLys 68
 Db 491 ---AAAGGAGATCAAAAGAACCAACCAAGGAATCCG-----ACCTCAA 450
 QY 68 sGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAl 88
 Db 449 AGTGATATTAAGTATATTTCCCTAATAGGCG---GGCCCAACCTATAGGATTATGATTA 393
 QY 88 aLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLys-----ThrHisLe 106
 Db 392 TATTAGCTTATATGGCAACTGGGATTTATCTAGGATTATACAAATTTTCCACTCACTT 333
 QY 106 uSerAlaSerIleProAspThrAspArg-----MetTrpTyrSerIleG1 121
 Db 332 TAGTACTGGAGATACCGCAAGAGATGAAGCAAGCAAAATACCTTTAGTTGGGAGCGG 273
 QY 121 yAlaThrTyrLysPheThrProAsnLeuSerValGlyPheAlaHisLeuArg-- 140
 Db 272 TTTTATTTAATTAATCCCAACCACTTCAATTTACCGCGGTTATGACAGGTAGCCGCTT 213
 QY 141 -----GlyLysLysLysHisPheValGluThrGlnAsnIleLysGlyLeu 156
 Db 212 TAGTATTAAAGACGGAACGACGAAAGATCATTTATCTACT----- 173
 QY 156 uLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGlyLeuAsnLeuAs 176
 Db 172 -----AACGGTTTAAATATACTGTCGG 150
 QY 176 nTyrArgPhe 179
 Db 149 CTATCGTTTC 140

```

RESULT 8
BQ706806          941 bp      mRNA      linear      EST 16-JUL-2002
LOCUS             AGENCOURT_8418322 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6281413
DEFINITION        5', mRNA sequence.
ACCESSION         BQ706806
VERSION           BQ706806.1 GI:21845705
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Dr. Mark Watson
                  cDNA Library Preparation: Rubin Laboratory
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LCM2474 row: c column: 14
                  High quality sequence stop: 621.
FEATURES          source
                  1..941
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6281413"
                  /lab_host="NIH_MGC_113"
                  /note="Organ: spleen; Vector: pONB7; Site:1: XhoI; Site:2:
                  EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGACAGAG(G). Library constructed by Ling Hong in the
                  Laboratory of Gerald M. Rubin (University of California,
                  Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                  Superscript II RT (Life Technologies). Note: this is a
                  NIH_MGC Library."
BASE COUNT        219 a 220 c 281 g 220 t 1 others
ORIGIN
Alignment Scores:
Pred. No.:        4.13      Length:      941
Score:            85.00     Matches:    42
Percent Similarity: 38.73%   Conserv:    25
Best Local Similarity: 24.28% Mismatches: 65
Query Match:      8.88%     Indels:    43
DB:              14        Gaps:      8

US-09-995-493-52 (1-179) x BQ706806 (1-941)
Qy 3 AsnGlyValLeuGlyProTyrIleGlySerLeuThrLeuLysLeuProAlaTyr 22
Db 382 AATGGGGACCTGCTCCTATGTGGGCAAGCTGGC---ACTGTACCTTCCT----- 432
Qy 23 TrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyr----- 39
Db 433 -----AGGAACCTCCACTACCCCTCCCTGATGGTTATAGTCGCCACTATGAAGATGGT 483
Qy 40 -----SerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLys 55
Db 484 TATCCAGGTGCAGTGATACTATGGCAGTCTGTCCGGGTGACCCCGCATGTAGGACGG 543
Qy 56 TyrGlnAspGly---SerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsn 74
Db 544 TATAGGCCAGCATGGGAAGGCTACCGGCACCTAGTAGACAGGATGTGTATGGGCCCAA 603
Qy 75 SerArgPheAlaIleGlyThrThr-----Tyr 83
Db 75 SerArgPheAlaIleGlyThrThr-----Tyr 83

```

```

Db 604 CCCAGGTTCCGGTAGGTGGGAGCAGCGTGGATCTGCATCTTCATCCAGAGCCTTAT 663
Qy 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla-----Ala 101
Db 664 GGGCTAGAGGATGAC-----CAGCGTAGTAGGGCTATGATGACCTGGATTATGGT 714
Qy 102 SerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGly 121
Db 715 ATGATGCTGATTATGGCACTGCCCGTTCGCAGCTGGGACACCCCTCGACCTCGTC-GGC 773
Qy 122 AlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArgGly 141
Db 774 GCCTCA-----GGTAAGGCAGCA-ATAAGGGGA 799
Qy 142 LysLysLysHisPheValGluThrGlnAsnIleLysGly 154
Db 800 AGAGAAAGGCTCTTCCCAAGACCAAGACAAGGGGG 838

RESULT 9
B1547625          672 bp      mRNA      linear      EST 05-SEP-2001
LOCUS             603191761F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5263190 5',
DEFINITION        mRNA sequence.
ACCESSION         B1547625
VERSION           B1547625.1 GI:15434937
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
                  cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
                  Toshiyuki and Piero Carninci (RIKEN)
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM11663 row: i column: 15
                  High quality sequence stop: 672.
FEATURES          source
                  1..672
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5263190"
                  /clone_lib="NIH_MGC_95"
                  /tissue_type="hippocampus"
                  /lab_host="DH10B"
                  /note="Organ: brain; Vector: pBluescriptR (modified
                  pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
                  ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
                  size-selected for average insert size 2.5 kb and
                  normalized to ROT 5. This is a primary library enriched
                  for full-length clones and constructed using the
                  Cap-trapper method (Carninci, in preparation). Library
                  constructed by M. Brownstein (NIH/NHGRI, National
                  Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT        146 a 172 c 209 g 145 t
ORIGIN
Alignment Scores:
Pred. No.:        3.3      Length:      672
Score:            84.00     Matches:    35
Percent Similarity: 40.16%   Conserv:    14
Best Local Similarity: 28.69% Mismatches: 53
Query Match:      8.78%     Indels:    20
DB:              13        Gaps:      6

```



```

RESULT 13
BH878604/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 600)
Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cs.hl.org
Plate: hs83 row: b column: 03
Seq primer: -21M3UnlvFwd
Class: shotgun
High quality sequence stop: 600.
Location/Qualifiers
1..600
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hs83b03"
/clone_lib="WGS-2maysF (JM107 adapted methyl filtered)"
/lab_host="JM107 or DH5a"
/note="Organ: Immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was neutralized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(-x/y reads in m13mp19, b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."
BASE COUNT 174 a 118 c 109 g 199 t
ORIGIN

Alignment Scores:
Pred. No.: 4,32 Length: 600
Score: 82,50 Matches: 31
Percent Similarity: 45,61% Conservativeness: 21
Best Local Similarity: 27,19% Mismatches: 53
Query Match: 8,62% Indels: 9
DB: 17 Gaps: 4

US-09-995-493-52 (1-179) x BH878604 (1-600)

QY 50 LysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
::: ||| ||::: ||| |||
DB 446 GAATCACTTGATGACAACTGAAGTCCCATCAGATAATGGCTCATTTGTTCTGATAAC 387
QY 70 GluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeu 89
||| :::: ||||| |||
DB 386 TGGGTTTCGATAGACACATCTAATCAGTTGGTACTAAAGGTTATGTGTCATCTGTGTAA 327
QY 90 ThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSer 109
::: |||::: :::: |||||::: |||
DB 326 GTCATCAGGTCC-----CTGCATTCATTAAACAACATCATACACCAAA 282
QY 110 IleProAspThrAspArgMetTrpTyrSerIleGlyAlaThrTyrLysPheThrProAsn 129

```

```

Db 281 TGTCTCTACAATA-----TACATGTACAGCTGGGGTGTTCATCTAAATTCACATGAAT 228
||| |||::: ||| :::: ||||| |||
QY 130 LeuSerValAsp--ValGlyPheAlaHisLeuArgGlyLysLysLysHisPheValGlu 148
::: |||::: ||| ||| ||| |||
Db 227 AAAATATATAGAAAAAAGTTTAAATCTAAAGAAAGGGCAAGCATTTGATCCAA 168
||| ||| ||| ||| ||| |||
QY 149 ThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyr 162
||| ||| ||| ||| ||| |||
Db 167 ---CAGACAGAAAGCTCTACTTGTATTGCATGTACCATTC 129
||| ||| ||| ||| ||| |||

RESULT 14
AW553447/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 649)
Tanaka,T.S., Jaradat,S.A., Llm,M.K., Kargul,G.J., Wang,X., Grahovac
,M.J., Pantano,S., Sano,Y., Plao,Y., Nagaraja,R., Dol,H., Wood,W.H.
III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
Plate: L0227 row: C column: 06
Seq primer: -21M13 Forward
High quality sequence stop: 649
POLYA=Yes.
Location/Qualifiers
1..649
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="plasmid:L0227C06-3"
/db_xref="taxon:10090"
/clone="L0227C06"
/clone_lib="NIA Mouse Newborn Ovary cDNA Library"
/sex="female"
/dev_stage="Newborn Ovary"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
7 Newborn Ovary. The double-stranded cDNA was synthesized
by Gibco's kit with an Oligo(dT) primer [NotI
primer-adaptor from Gibco/BRL]
15'-PGACTAGTCTAGATCGGAGCGCGCCCTTTTTTTT-3'] from
2.5µg of total RNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Glbco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Yulan Plao."
BASE COUNT 182 a 162 c 137 g 168 t

```

```

ORIGIN
Alignment Scores:
Pred. No.: 4.88 Length: 649
Score: 82.50 Matches: 40
Percent Similarity: 37.79% Conservative: 25
Best Local Similarity: 23.26% Mismatches: 60
Query Match: 8.62% Indels: 47
DB: 10 Gaps: 9

US-09-995-493-52 (1-179) x AW553447 (1-649)
Qy 5 ValLeuGlyProTyrIleGlyLysGly---SerLeuThrLeuLysLeuPro---AlaTyr 22
Db 590 CTGTTGCTTCCTATGTCAGCAGTGAGCTGAACATGATCATCATCTCCAGATGAGCAC 531
Qy 23 TrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSerTyrLys 42
Db 530 GTTGAACGTGAGT-----ACAGTGAAAGAGGAAGTAACCTACGAG---AAA 489
Qy 43 TyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyr 62
Db 488 TTTATAGAGTGCAAGCGCTGCACAAGATG-----GACGGAGAGAGGTA 444
Qy 63 GluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPheAlaIleGlyThrThr 82
Db 443 GAAGTATTCTCCCAAGATTAAAGCTGGAGAGAAT----- 408
Qy 83 TyrSerLeuAsnAspAlaLeu----- 89
Db 407 TATAACATGAACGATGCCCTCTACAAAGTTGGCGATGACTGATGCTTTGGCGGAGGGCA 348
Qy 90 -----ThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThr 104
Db 347 GACTTTCTGGAAGTCTTCCCAAGCAAGGCTGTTCTGTCTAAGTTGTGCAATAGGCC 288
Qy 105 HisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGly----- 121
Db 287 TTTGTGGAGGTTAATGAGGAGGCGCAGAGCTGCACGCTGCATACAGCTGGCATGACG 228
Qy 122 ---AlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHisLeuArg 140
Db 227 GTGAGGTGCATGACATTCACTCCCGCTTCTGTCGCCACCAACCCCTTCCTTCCTTC 171
Qy 141 GlyLysLysLysHisPheValGluThrGlnAsnIle 152
Db 170 -----ATTCAACATGTTAAGACCAATGGAATT 144

RESULT 15
BJ041818 683 bp mRNA linear EST 07-DEC-2001
LOCUS
DEFINITION
BJ041818 NIBB Mochii normalized xenopus neurula library xenopus
laevis cDNA clone XL027106 5', mRNA sequence.
ACCESSION
BJ041818
VERSION
BJ041818.1 GI:17417364
KEYWORDS
EST.
SOURCE
African clawed frog.
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 683)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara
Y.
TITLE
Expressed genes in x. laevis embryo
JOURNAL
Unpublished (2001)
COMMENT
Contact: Tadasu Shin-I
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin1@genes.nig.ac.jp.

FEATURES
source
Location/Qualifiers
1..683
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL027106"
/clone_lib="NIBB Mochii normalized Xenopus neurula
library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 176 a 171 c 167 g 167 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 5.29 Length: 683
Score: 82.50 Matches: 35
Percent Similarity: 40.58% Conservative: 21
Best Local Similarity: 25.36% Mismatches: 57
Query Match: 8.62% Indels: 25
DB: 13 Gaps: 6

US-09-995-493-52 (1-179) x BJ041818 (1-683)
Qy 4 GlyValLeuGlyProTyrIleGlyLysGlySerLeuThrLeuLysLeuProAlaTyrTrp 23
Db 242 GGTATCGTGTGCTGCTAT---GGTGCAGCTGACCGTACCAACCTGCCAAGAGCT----- 292
Qy 24 GluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla----- 36
Db 293 CAACCTCTTGCAACGCTAAAAAGCTGACAGTGGCTACTGCTCTGAAGTACGACGCG 352
Qy 37 -----IleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGly 54
Db 353 AACACATCTACCTGGCAGCGAAGTACGGTGAAACCCGTAACGCTACGCCGATCCTAAT 412
Qy 55 LysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsn 74
Db 413 AAATTTACAACACCACCGGCTTCGCCACAAACGCAAGAGCTTCTGTAGTTGCCCAA 472
Qy 75 SerArgPheAlaIleGlyThrTyrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGly 94
Db 473 TACCAGTTCCGATTTTCGGT-----CTGCGTCCGTC 502
Qy 95 LeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAspThrAsp 114
Db 503 ATCCCTTACACCAATCTAAGCGAAAGAC-----GTAGAAGGTATCGGTGATGTTGAT 556
Qy 115 ArgMet---TrpTyrSerIleGlyAlaThrTyrLysPheThrProAsnLeuSer 131
Db 557 CTGCTGAACCTACTTTGAAGTGGCGGCAACCTACTACTTCAACAAACATGCTCC 610

RESULT 16
BM930833 744 bp mRNA linear EST 13-MAR-2002
LOCUS
DEFINITION
BM930833 UI-E-EJ1-ajb-k-10-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
UI-E-EJ1-ajb-k-10-0-UI 5', mRNA sequence.
ACCESSION
BM930833
VERSION
BM930833.1 GI:19390006
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 744)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477

```

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation by: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 REVERSE.

FEATURES

source

Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UI-E-EJ1-a1b-k-10-0-UI"
/clone_lib="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT; optic nerve, CCATTAGTGC; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT

ORIGIN

220 a 99 c 153 g 271 t
Alignment Scores:
Pred. No.: 6.03 Length: 744
Score: 82.50 Matches: 25
Percent Similarity: 43.3% Conservative: 14
Best Local Similarity: 27.78% Mismatches: 30
Query Match: 8.62% Indels: 21
DB: 14 Gaps: 3

US-09-995-493-52 (1-179) x BW930833 (1-744)

Qy 84 SerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaLaSerLys 103
Db 739 TCATTAATAAATAATACAGTACTTTGGGTAGGA-----AAAGAAGTGCATTAA 686
Qy 104 ThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 685 CTTCAATTATGTTTTCCTCCACAAAGAGCGGCCACAATTT---GTACATCCACCA 629
Qy 124 TyrLysPheThrProAsnValAspValGlyPheAlaHisLeuArgGlyLysLys 143
Db 628 TACAATACACAAATTCATTATCTTCAAAATACTATATACACAAAGCGGAGAA 569
Qy 144 LysHisPheValGluThrGlnAsnIleLysGlyLeuLeuValGluAlaAspTyrThr 163
Db 568 -----TATACA 563
Qy 164 ThrLysAlaThrAlaAsnLeuTyrGlyLeu 173
||| |||:||||: |||

Db 562 GCATGCTCTTTTGGTAACATATGAATTG 533

RESULT 17

AZ573577

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ573577 665 bp DNA linear GSS 15-MAY-2001
319PVA06 Pv MBN #30 Plasmodium vivax genomic 3', DNA sequence.
AZ573577
GSS
malaria parasite P. vivax.
Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 665)
Carlton, J.M.-R. and P. berghei gene sequence tag projects
The Plasmodium vivax and P. berghei gene sequence tag projects
Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.

FEATURES

source

Location/Qualifiers

1..665
/organism="Plasmodium vivax"
/strain="Salvador I (Collins, W. 1972. J. Parasitol. 69, 497-598)"
/db_xref="taxon:5855"
/clone_lib="Pv MBN #30"
/dev_stage="asexual blood forms"
/lab_host="Salmirli bollivensis"
/note="Vector: pBluescript SK(+) vector DNA, phagemid excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V; Host leukocytes were extracted from P. vivax infected blood using the following methods: first, infected blood was activated by the addition of 0.5 ml of ADP (40mg/ml) per 10 ml blood. Then blood was passed over a column of acid washed 0.1 mm glass beads, then through a Plasmidipur filter, followed by passage through a column of pre-wet Whatman Cfil powder (1:2 ratio volume of blood to Cfil), and finally centrifuged through a 50% Percoll density cushion. Purified DNA was digested with mung bean nuclease in the presence of 4% formamide at 50°C as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). Digested DNA was blunt-ended using T4 DNA Polymerase and size fractionated over a Sepharose CL-2B column. Fractions in the size range 500bp-4kb were ligated into the EcoRV site of pBluescript SK(+), and E. coli XL-10 Gold transformed with the ligation mixture."

BASE COUNT

ORIGIN

205 a 168 g 134 t

Alignment Scores:

Pred. No.: 5.88 Length: 665
Score: 82.00 Matches: 35
Percent Similarity: 39.46% Conservative: 23
Best Local Similarity: 23.81% Mismatches: 49
Query Match: 8.57% Indels: 40
DB: 17 Gaps: 6

US-09-995-493-52 (1-179) x AZ573577 (1-665)

Qy 50 LysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
||| |||:||||: |||
Db 39 AAAGAACGTCAGATGTGAACCGAGATGAAGCAGCAGCTCCGCTTCCCAATATCAAG 98
Qy 70 -----GluTyrLysAspAsnSerArgPheAla---IleGly 80
||| |||:||||: |||
Db 99 TACGGCAATATTTCTTATGCTATGAGGTGAGGAACCAAGATTAAGTTTTCAGACCTCTAC 158

```

Qy 81 ThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAla 100
   :::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 AGCGTTTACACCAAAATGTCAA- - - - - CAGAGGATCAGCG 197
Qy 101 AlaSerLysThrHisLeuSerAlaSerIleProAspThrAspArgMetTrpTyrSerIle 120
   :::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 CACACGAAACGCTGCTCAAAATCGAACATCCCAATTTGGACAAGTTTCATAAATAGCAGC 257
Qy 121 GlyAlaThrTyrLysPheThrPro- - - - - AsnLeuSerVal 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 AGGAATAGGAGCTCTTCCCGCCGAGGAGAGATGCCGGCGGGAACACCGGTGCA 317
Qy 133 AspValGlyPheAlaHisLeuArgGlyLysLysHisPheValGluThrGlnAsnIle 152
   :::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 AATATTGGCGGCCCAAAATATTGGCGGC- - - - - GCAAATGTT 353
Qy 153 LysGlyLeuLeuLeuValGluAlaAspTyrThrThrLysAlaThrAlaAsnLeuTyrGly 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 AGCGCGCTAAATCTC- - - - - AGCGCGTGAATGCTGGCGGC 389
Qy 173 LeuAsnLeuAsnTyrArgPhe 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 CTAACACTACAATTACAACTAC 410

RESULT 18
BM396839 725 bp mRNA linear EST 17-JAN-2002
LOCUS 5009-0-25-G03.t.2 Chilcoat/Turkewitz cDNA (large fraction)
DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION BM396839.1 GI:18196892
VERSION 1
KEYWORDS EST.
SOURCE Tetrahymena thermophila.
ORGANISM Tetrahymena thermophila.
REFERENCE Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
AUTHORS Hymenostomatida; Tetrahymenina; Tetrahymena.
TITLE Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel
J. and Klobutcher,L.
JOURNAL EST from Tetrahymena thermophila, strain CU428.1, growing cells
COMMENT Unpublished (2002)
Contact: Turkewitz AP
Molecular Genetics and Cell Biology
University of Chicago
920 E. 58th Street, Chicago, IL 60637, USA
Tel: 773 702 4374
Fax: 773 702 3172
Email: apturkew@midway.uchicago.edu
Seq primer: f3.

FEATURES
source Location/Qualifiers
1..725
/organism="Tetrahymena thermophila"
/strain="CU428.1"
/db_xref="taxon:5911"
/clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
/notes="vector: Bluescript SK+; Details on library
preparation can be found in Chilcoat and Turkewitz (2001)
Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT 217 a 129 c 149 g 230 t
ORIGIN

Alignment Scores:
Pred. No.: 6.73 Length: 725
Score: 82.00 Matches: 39
Percent Similarity: 29.63% Conservative: 17
Best Local Similarity: 20.63% Mismatches: 49
Query Match: 8.57% Indels: 84
DB: 13 Gaps: 9

US-09-995-493-52 (1-179) x BM396839 (1-725)

Qy 21 AlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAlaIleHisTyrSer 40

```

```

Db 182 GCTTCTCGGGTATTAGGGTTATGAAGAAATTGTTACCGAATTTGGTACCCACAAGGT 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 41 TyrLysTyrThrGluTrpSerArgPheLysGluLeuArgGlyLysTyrGlnAspGlySer 60
   :::::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 CATAATTAT- - - - - TGGCCCCAAATTTTCATGCTTG- - - GGTACCTACGATTCGGTTCA 292
Qy 61 - - - - - GlyTyrGluAlaPheThrLysLysGluGluTyrLysAspAsnSerArgPhe 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 GTCAGACAGGTTTCCCAAGTTTTCGCCAGA- - - - - AATTGCGGTAAAC 334
Qy 78 AlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyr 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 TGTCATGTTATGATTTACAGAGTAGCATTTACTTA- - - - - TTG 373
Qy 98 AspLysAlaAlaSerLysThrHisLeuSerAlaSerIleProAsp- - - - - 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 GATAAGGCTTACAGATAATTAGAATTAGCTCAAAATGGTTCGCGATTTCATATTCAACCT 433
Qy 113 - - - - - - - - - - - - - - - - - - - - - ThrAspArg- - - - - 115
Db 434 GCCCATCAACATTTTAAACAATATTACTATCAAGAAATGGGATGAGAGATAGATCAT 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 115 - - - - - - - - - - - - - - - - - - - - - 115
Db 494 TGCATCACATCTATCTCTCTTACTTCTCTTAGGATCAAGCTAAGAAATGCTAAGCGTGT 553
Qy 116 MetTrp- - - - - - - - - - - - - - - - - - - - - Tyr 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 GTCGCGCTACTGATTCTCTAAGATTAAAGCTTAGACCTGGTGTATTAACTACATTTAT 613
Qy 119 SerIleGlyAlaThrTyrLysPheThrProAsnLeuSerValAspValGlyPheAlaHis 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 AACATTCTACTGCTTATCACTTCTCACTCTCTTTCGTTATGGATGTT- - - - - 661
Qy 139 LeuArgGlyLysLysHisPheVal 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 - - - - - CCTAAGGCGCAAAATACTTCATC 682

RESULT 19
AW788126 386 bp mRNA linear EST 01-MAY-2001
LOCUS C00212-F Lambda Zap, Stratagene Blumeria graminis f. sp. hordei
DEFINITION cDNA clone C00212, mRNA sequence.
ACCESSION AW788126
VERSION 1
KEYWORDS AW788126.1 GI:13899723
SOURCE Blumeria graminis f. sp. hordei.
ORGANISM Blumeria graminis f. sp. hordei.
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
AUTHORS Erysiphales; Erysiphaceae; Blumeria.
Thomas,S.W., Rasmussen,S.W., Glaring,M.A., Rouster,J.A. and Oliver
R.P.
TITLE Gene identification in the fungal pathogen Blumeria graminis by
expressed sequence tag analysis
JOURNAL Unpublished (2000)
COMMENT Contact: Rasmussen,S.W.
Department of Yeast Genetics
Carlsberg Laboratory
10 GL. Carlsbergvej, DK-2500, Copenhagen, Denmark
Tel: 45 3327 5230
Fax: 45 3327 4766
Email: swr@crc.dk
High quality sequence stop: 386
POLYA-No. Location/Qualifiers
1..386
/organism="Blumeria graminis f. sp. hordei"
/db_xref="taxon:62688"
/clone="C00212"
/cell_type="conidia"

```

/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-3624"
/clone_lib="Gm-cl027"
/tissue_type="cotyledons of 3- and 7-day-old Williams seedlings"
/lab_host="DHIOB"
/note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 3- and 7-day-old Williams seedlings which were propagated on paper towels with distilled water. The cotyledons were flash-frozen in liquid nitrogen, then lyophilized for 72 hours. Unequal amounts of mRNA was used for cDNA synthesis. Stratagene's cDNA Synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGACTGTTCGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The xhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all xhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript(tm) II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Kelm and Dr. Virginia Corvelli."

BASE COUNT 125 a 76 c 92 g 124 t
ORIGIN

Alignment Scores:
Pred. No.: 3,31 Length: 417
Score: 81.50 Matches: 26
Percent Similarity: 45.24% Conservative: 12
Best Local Similarity: 30.95% Mismatches: 25
Query Match: 8,52% Indels: 21
DB: 10 Gaps: 4

US-09-995-493-52 (1-179) x AW759240 (1-417)
QY 2 HisAsnGlyValLeuGlyProTyrTle-----GlyLysGlySerLeuThr 16
Db 128 TATTCGGCTTTGTAACGACAAATTTCATGTTGCTGGGAGCAGGAAGC----- 181
QY 17 LeuLysLeuProAlaTyrTrpGluLeuSerGlyPheHisGlnLeuThrAspGlnTrpAla 36
Db 182 -----CACTTGCAAAATTTTCAGCCAAGTAACCCCTAAGTGAGT 220
QY 37 IleHisTyrSerTyr-----LysTyrThrGluTrpSerArgPheLysGlu 51
Db 221 CTTTACAGAGACTACCATTTTGGGTTTGTCAAGTTGCAGCAGCATTCATCTCTC 280
QY 52 Leu-----ArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
Db 281 CTCCTTTGAATACAAGAAAAGCAGCTGATCGAANGGTCTACGATCTCTTTTACCATTTCAGG 340
QY 70 GluTyrLysAsp 73
Db 341 GATTACAAAGAT 352

RESULT 21
BM178271 563 bp mRNA linear EST 06-DEC-2001
LOCUS

```

sa71a04.y1 Gm-cl072 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl072-5000 5' similar to TR:Q9SW37 Q9SW37 HYPOTHETICAL 61.4 KD
PROTEIN. ;, mRNA sequence.
BML178271
BML78271.1 GI:17401489
EST.
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 563)
Shoemaker,R., Kelm,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Contact: Shoemaker R/Public Soybean EST Project
Unpublished (1999)
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -4ORP from Gibco
High quality sequence stop: 422.
Location/Qualifiers
1..563
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl072-5000"
/contig_lib="Gm-cl072"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/notes="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar PI
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E.coli Electromax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Steve Clough (Lila Vodkin lab,
University of Illinois)."
174 a 99 c 102 g 188 t
BASE COUNT
ORIGIN
Alignment Scores:
red. No.: 5.27 Length: 563

```



```

source      1. .659
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODM002YF06"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 165 a 220 c 199 g 75 t
ORIGIN

Alignment Scores:
Pred. No.: 7.81 Length: 659
Score: 81.00 Matches: 41
Percent Similarity: 41.01% Conservatives: 16
Best Local Similarity: 29.50% Mismatches: 44
Query Match: 8.46% Indels: 38
DB: 12 Gaps: 9

US-09-995-493-52 (1-179) x BG740114 (1-659)

Qy 38 HlsTyrSerTyrLys-----TyrThrGluTrpSerArgPhe 49
||||| |||
Db 97 CACTACAGCGTCAAGGCTCAGTTCAGACGTCCTACACGGCGGGCGGCTGTC 156
||||| |||

Qy 50 LysGluLeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 69
||||| |||
Db 157 AAGGAG-----AGCTGGCGGGCGGGCGGCGGAGGCGACACGGCGGC--- 204
||||| |||

Qy 70 GluTyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAlaLeu 89
||||| |||
Db 205 -----CTGACGCGCTCGCATCGCAGATCTCGGAGCTCGCTCGCTGGGC----- 249
||||| |||

Qy 90 ThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHlsLeuSerAlaSer 109
||||| |||
Db 250 -----CGCGAGGGC---GCCACAGCAAGGCGGGCGGCGCT-----GCCGCC 288
||||| |||

Qy 110 IleProaspThrAspArgMetTrpTyrSerIleGly---AlaThrTyrLys----- 125
||||| |||
Db 289 GTCTCAGACGAGCTCGCGCAGGTGACGAGTCCACCGCTCCGACCAAGGACGACAGC 348
||||| |||

Qy 126 -----PheThrProAsnLeuSerValAspValGlyPhe----- 136
||||| |||
Db 349 CGCCTGTGCGACACAGCTCCACCTCTCGGACAGCGGCTCGCAAGTACAGCACTCCTC 408
||||| |||

Qy 137 ---AlaHlsLeuArgGlyLysLysLysLysPheValGluThrGlnAsnIleLysGly 154
||||| |||
Db 409 CCAGAGCACCTTCGTGGCGCACAGCAGGGTCACAGGATGCCCGCAGATGTGCAAGGG 465
||||| |||

RESULT 24
LOCUS AL531424 LTI_NFL001_NBC4 Homo sapiens cDNA clone CSODM002YF06 5
DEFINITION prime, mRNA sequence.
ACCESSION AL531424
VERSION AL531424.1 GI:12794917
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 873)
JOURNAL Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .873

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODM002YF06"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 187 a 221 c 254 g 209 t
ORIGIN

Alignment Scores:
Pred. No.: 12.1 Length: 873
Score: 81.00 Matches: 34
Percent Similarity: 40.16% Conservatives: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.46% Indels: 20
DB: 9 Gaps: 6

US-09-995-493-52 (1-179) x AL531424 (1-873)

Qy 32 ThrAspGlnTrpAlaIleHlsTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
||||| |||
Db 98 ACTGTGAGATGGTGTGCAGTGTCCGAGCATGAGGCCACCTAAGTGCACAGATTTCGCGC 157
||||| |||

Qy 52 ---LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu 70
||||| |||
Db 158 CATATGAAGAAGCGTCATTCCATCCGATGTCCTGCTGTGTGAAGAAATCCTCC 217
||||| |||

Qy 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAlaLeuThr 90
||||| |||
Db 218 TACCTTGATTCATCAGG---GCCATTGCGCAAGGAGCG-----GATGCTGTGACA 268
||||| |||

Qy 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHlsLeuSer----- 107
||||| |||
Db 269 CTGGATGAGGTTGGTGTATGATGCTTACCTGGCTCCCAATAACCTGAAGCTGTGGTG 328
||||| |||

Qy 108 -----AlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
||||| |||
Db 329 GCAGAGTTCATGGGTCAAAAGAGGATCCACAGACTTCTATTATGCTGTGCTGTGGTG 388
||||| |||

Qy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HlsLeuArgGly 141
||||| |||
Db 389 AAGAAG-----CATAGTGGCTTCCAGATGAACCAAGCTTCGAGCG 427
||||| |||

RESULT 25
LOCUS W64749 484 bp mRNA linear EST 10-JUN-1996
DEFINITION md88c04.1 Soares mouse embryo NDMEJ3.5 14.5 Mus musculus cDNA
clone IMAGE:385062 5' similar to PIR:A47411 A47411
ADPribosylarginine hydrolase ; mRNA sequence.
ACCESSION W64749
VERSION W64749.1 GI:1372452
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 484)
REFERENCE Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
AUTHORS

```


Gelsei,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wyllie,T., Lennon,G., Soares,B., Willson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:236894
 Possible reversed clone: similarity on wrong strand
 Seq primer: ETPrimer
 High quality sequence stop: 342.
 Location/Qualifiers
 1..484
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:385062"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGAGCGCGCGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 BASE COUNT 117 a 121 c 141 g 105 t;
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.62 Length: 484
 Score: 80.50 Matches: 29
 Percent Similarity: 40.57% Conservatives: 14
 Best Local Similarity: 27.36% Mismatches: 42
 Query Match: 8.41% Indels: 21
 DB: 14 Gaps: 5
 US-09-995-493-52 (1-179) x W64749 (1-484)
 QY 38 HisTyrSerTyrLysTyrThrGluTyrSerArgPheLysGluLeuArgGlyLysTyrGln 57
 |||||:||||| ||||| :|||: |||||:|||||
 Db 68 CACTGCTCTACTTCGAGAAAGATGGAAAGTACCTGGCTTAGAGGA---ATTGTG 124
 QY 58 AspGlySerGlyTyrGluAlaPheThrLysLysGluGlyTyrLysAspAsnSerArgPhe 77
 |||||:||||| ||||| :|||: |||||:|||||
 Db 125 GACGGCAACTCGGCTCCCGCTTCCTCCGACGCCCTTTGGTGTGAAGGAAGGATCAGTTTC 184
 QY 78 AlaIleGlyThrThrTyrSer----- 84
 ||| :|||:|||||
 Db 185 TACATCGAGCTGAGCTACTCGGGCTGGGTGGCAGCAGCGGACACCATGCCCCATGATT 244
 QY 85 LeuAsnAspAlaLeuThrLeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThr 104
 |||||:||||| ||||| :|||: |||||:|||||
 Db 245 GCCTACGATGCCCTC---CTGGCCGGGGGGATTCTCGAAGGAGCTCGCACACAGAGCC 301
 QY 105 HisLeuSerAlaSerIleProAspThrAspArg-----MetTrpTyrSerIle 120
 |||||:|||||

Db 302 TTCTTCCACTGTGGAGACACTGATTCACGGCGCCCATTCCTGGCTGCTGGTGB----- 355
 OY 121 GlyAlaThrTyrLysPhe 126
 ||| ||| |||
 Db 356 GGAGTTATGTACGGCTTT 373
 RESULT 26
 W64824
 LOCUS
 DEFINITION
 W64824 484 bp mRNA linear EST 10-JUN-1996
 md88g04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:385110 5, similar to PIR:A47411 A47411
 ADPribosylarginine hydrolase ;, mRNA sequence.
 W64824
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
 1 (bases 1 to 484)
 REFERENCE
 AUTHORS
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geiseli,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wyllie,T., Lennon,G., Soares,B., Willson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:236942
 Possible reversed clone: similarity on wrong strand
 Seq primer: ETPrimer
 High quality sequence stop: 343.
 Location/Qualifiers
 1..484
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:385110"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGAGCGCGCGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 BASE COUNT 117 a 122 c 143 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.62 Length: 484
 Score: 80.50 Matches: 29
 Percent Similarity: 40.57% Conservatives: 14
 Best Local Similarity: 27.36% Mismatches: 42
 Query Match: 8.41% Indels: 21
 DB: 14 Gaps: 5

REFERENCE AUTHORS	TITLE JOURNAL COMMENT	FEATURES source	BASE COUNT
1 (bases 1 to 599)	Endocrine Pancreas Consortium Unpublished (2000) Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@biohp.harvard.edu Library was constructed by Dr. Hiroshi Inoue DNA sequencing on Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@imgate.wustl.edu) High quality sequence stop: 444. Location/Qualifiers 1. .599 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Human Fetal Pancreas 1A" /tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20 weeks, StrataGene #738023)" /dev_stage="Fetal Pancreas" /note="Vector: pBluescript SK(-); Site1: NotI; Site2: XhoI; cDNA made by oligo-dT priming. Size-selected by CHROMA SPIN-1000 (Clontech). Average insert size ~1kb. 5', XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel: 314-362-1916, Fax:314-747-2692."	124 a 150 c 180 q 145 t	

Db	375	AAGAAG-----				GATAGTGGCTTCACAGTAACACCAGCTTCGAGGC	413
Qy	142	Lyslys 143					
Db	414	AAGAAG 419					
RESULT#	31						
LOCUS	AL531052						
DEFINITION	AL531052 LTI_NFL001_NBC4 Homo sapiens cDNA clone CSODM005VC21 5 prime, mRNA sequence.	611 bp	mRNA	llinear	EST l3-FEB-2001		
ACCESSION	AL531052						
VERSION	AL531052.1	GI:12794545					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1. (bases 1 to 611)						
TITLE	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.						
JOURNAL	Full-length cDNA libraries and normalization						
COMMENT	Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.						
FEATURES	Email: Location/Qualifiers						
source	1. .611						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
	/clone="CSODM005VC21"						
	/clone_lib="LTI_NFL001_NBC4"						
	/sex="male"						
	/tissue_type="neuroblastoma cells"						
	/lab_host="DH10B"						
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"						
BASE COUNT	125 a	154 c	184 g	147 t	1 others		
ORIGIN							
Alignment Scores:							
Pred. No.:	9.35					Length:	611
Score:	80.00					Matches:	34
Percent similarity:	40.16%					Conservative:	15
Best Local Similarity:	27.87%					Mismatches:	53
Query Match:	8.36%					Indels:	20
DB:	9					Gaps:	6
US-09-995-493-52 (1-179) x AL531052 (1-611)							
Qy	32	ThrAspGlnTrpAlaIleHisTySerTyrLysTyrThrGluTrpSerArgPheLysGlu	51				
Db	90	ACTGTGAGATGGTGCGAGTCCGAGCATGCGCCACTAAGTCCCAGAGTTCCCGCAC	149				
Qy	52	--LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGlu	70				
Db	150	CATATGAACCGCTCATTCCATCCGATGGTCCCACTGTTGCTGTGTCAAGAACCTCC	209				
Qy	71	TyrLysAspAsnSerArgPheAlaIleGlyThrTyrSerLeuAsnAspAlaLeuThr	90				
Db	210	TACCTTGATTCGATCAGG---GCCATTGCGCAACGAAGCG-----GATGCTGTGACA	260				
Qy	91	LeuArgAlaGlyLeuAlaTyAspLysAlbaSerLysThrHisLeuSer-----	107				

```

Db 261 CTGGATGCAGGTTTGGTGTATGATGCTTACCTGGCTGCCAATAAAGCTGAAGCCTGTGGTG 320
Qy 108 -----AlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 321 GCAGAGTCTTATGGGTCAAAAGAGATCCACAGACTTCTATTATGCTGTGCTGTGGTG 380
Qy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 381 AAGAAG-----GATAGTGCTTCCAGATGAACACCTTCGAGGC 419
Qy 142 LysLys 143
Db 420 AAGAAG 425

RESULT 32
LOCUS BG700910 615 bp mRNA linear EST 07-MAY-2001
DEFINITION 602681855F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814772 5',
mRNA sequence.
ACCESSION BG700910
VERSION BG700910.1 GI:13970725
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0710 row: m column: 13
High quality sequence stop: 612.
FEATURES
Location/Qualifiers
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4814772"
/clone_lib="NIH_MGC_95"
/tissue_type="hnpocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTGTGTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to RQ7 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 129 a 156 c 185 g 145 t
ORIGIN

Alignment Scores:
Pred. No.: 9.44 Length: 615
Score: 80.00 Matches: 34
Percent Similarity: 40.16% Conservative: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.36% Indels: 20
DB: 12 Gaps: 6

US-09-995-493-52 (1-179) x BG700910 (1-615)
Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51

```

```

Db 121 ACTGTGAGATGCTGTCAGTGTGGAGCATGAGGCACCTAAGTGCAGAGTTTCCGCGAC 180
Qy 52 ---LeuArgGlyLysTyrClnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 181 CATATGAAGAAGCGTCATTCATCCGATGTCGCCAGTGTGGTGTGTAAGAAGACCTCC 240
Qy 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThr 90
Db 241 TACCTTGATGTCATCAGG---GCCATTGGCGCAACGACGCG-----CATGCTGTGACA 291
Qy 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaLaSerLysThrHisLeuSer-----107
Db 292 CTGGATGCAGGTTTGGTGTATGATGCTTACCTGGCTCCCAATAACCTGAAGCCTGTGGTG 351
Qy 108 -----AlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 352 GCAGAGTCTTATGGGTCAAAAGAGATCCACAGACTTCTATTATGCTGTGCTGTGGTG 411
Qy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 412 AAGAAG-----GATAGTGCTTCCAGATGAACACGCTTCGAGGC 450
Qy 142 LysLys 143
Db 451 AAGAAG 456

RESULT 33
LOCUS BG639058 617 bp mRNA linear EST 15-JUL-2002
DEFINITION hd30h04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone hd30h04 5', mRNA sequence.
ACCESSION BG639058
VERSION BG639058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 617)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
,J.W., Bouffard,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL Contact: Wistow G
COMMENT Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: grame@helix.nih.gov
Plate: 30 row: h column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..617
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="hd30h04"
/clone="hd30h04"
/tissue_type="Retina"
/dev_stage="Adult"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid system full
details of which are contained in the manufacturer's

```


647
34


```

Qy 108 -----AlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 333 GCAGAGTTCTATGGTCAAAAGAGATCCACAGACTTTCTATTATGCTGTGGTG 392
Qy 124 TyrLysPheThrProAsnLeuScrValAspValGlyPheAla-----HisLeuArgGly 141
Db 393 AAGAAG-----GATAGTGGCTTCCAGATGAACCAAGCTTCGAGGC 431
Qy 142 LysLys 143
Db 432 AAGAAG 437

RESULT 40
BE006001/c
LOCUS BE006001 665 bp mRNA linear EST 05-JUN-2000
DEFINITION RC0-BN0121-280300-032-b06 BN0121 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE006001
VERSION BE006001.1 GI:8266234
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 665)
AUTHORS Dias Ncto,E., Garcia Correa,R., Verjovski-Almeida,S.; Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC0-BN0121-280
300-032-b06&t3-2000-03-28&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 664.
FEATURES
source
1..665
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0121"
/dev_stage="Adult"
/notes="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent applicaion
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 169 a 188 c 159 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 10.7 Length: 665
Score: 80.00 Matches: 34
Percent Similarity: 40.16% Conservative: 15
Best Local Similarity: 27.87% Mismatches: 53
Query Match: 8.36% Indels: 20
DB: 10 Gaps: 6

```

```

US-09-995-493-52 (1-179) x BE006001 (1-665)
Qy 32 ThrAspGlnTrpAlaIleHisTyrSerTyrLysTyrThrGluTrpSerArgPheLysGlu 51
Db 646 ACTGTGAGATGGTGTGCGAGATGCGAGCCACTAAGTCCAGAGCTTCCGCGAC 587
Qy 52 ---LeuArgGlyLysTyrGlnAspGlySerGlyTyrGluAlaPheThrLysLysGluGlu 70
Db 586 CATATGAAGAAGCGTCATTCATCCATCCGATGGTCCAGTGTCTTGTGTGAAGAAAGCCTCC 527
Qy 71 TyrLysAspAsnSerArgPheAlaIleGlyThrThrTyrSerLeuAsnAspAlaLeuThr 90
Db 526 TACCTTGATTCATCAGG---GCCATTTCGCGCAACGAAGCG-----GATGCTGTGACA 476
Qy 91 LeuArgAlaGlyLeuAlaTyrAspLysAlaAlaSerLysThrHisLeuSer----- 107
Db 475 CTGGATCAGGTTTGTGTATGATGCTTACCTGCTCCCAATAACCTGAAGCCTGTGGTG 416
Qy 108 -----AlaSerIleProAspThrAspArgMetTrpTyrSerIleGlyAlaThr 123
Db 415 GCAGAGTTCTATGGTCAAAAGAGATCCACAGACTTTCTATTATGCTGTGGTG 356
Qy 124 TyrLysPheThrProAsnLeuSerValAspValGlyPheAla-----HisLeuArgGly 141
Db 355 AAGAAG-----GATAGTGGCTTCCAGATGAACCAAGCTTCGAGGC 317
Qy 142 LysLys 143
Db 316 AAGAAG 311

```

Search completed: May 19, 2003, 19:57:24
Job time : 1088 secs

GenCore version 5.1.4_p5_4578.
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 09:39:39 ; Search time 15 Seconds
(without alignments)
351.114 Million cell updates/sec

Title: US-09-995-493-52

Perfect score: 957

Sequence: 1 QHNGVLGYIGKSLTLKLP.....ADYTTKATNLYGLNLNRYF 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	559	58.4	455	2	US-08-472-172-4
2	549	57.4	459	2	US-08-472-172-2
3	502.5	52.5	432	2	US-08-472-172-6
4	252	26.3	77	2	US-08-472-172-12
5	129	13.5	34	1	US-08-475-989-8
6	129	13.5	34	2	US-08-475-985-8
7	129	13.5	34	2	US-08-472-172-20
8	129	13.5	34	3	US-08-256-839-8
9	111	11.6	32	1	US-08-475-989-10
10	111	11.6	32	2	US-08-475-985-10
11	111	11.6	32	2	US-08-472-172-22
12	111	11.6	32	3	US-08-256-839-10
13	89	9.3	459	2	US-08-810-655A-2
14	87.5	9.1	35	1	US-08-475-989-14
15	87.5	9.1	35	2	US-08-475-985-14
16	87.5	9.1	35	2	US-08-472-172-26
17	87.5	9.1	35	3	US-08-256-839-14
18	85.5	8.9	340	2	US-08-355-844-1
19	85.5	8.9	340	5	PCT-US95-16126-1
20	83.5	8.7	261	4	US-08-961-083-152
21	83.5	8.7	362	1	US-08-183-214-2
22	82	8.6	54	1	US-08-475-989-49
23	82	8.6	54	2	US-08-475-985-49
24	82	8.6	54	3	US-08-256-839-49
25	82	8.6	188	1	US-08-486-715-5
26	82	8.6	188	1	US-08-486-719-5
27	82	8.6	188	1	US-08-476-100-5

28 82 8.6 188 3 US-08-475-749-5 Sequence 5, Appl
29 81.5 8.5 362 1 US-08-183-214-10 Sequence 10, Appl
30 81 8.5 188 1 US-08-286-767-3 Sequence 3, Appl
31 80 8.4 38 1 US-08-475-989-13 Sequence 13, Appl
32 80 8.4 38 2 US-08-475-985-13 Sequence 13, Appl
33 80 8.4 38 3 US-08-472-172-25 Sequence 25, Appl
34 80 8.4 38 3 US-08-256-839-13 Sequence 13, Appl
35 80 8.4 53 1 US-08-475-985-48 Sequence 48, Appl
36 80 8.4 53 2 US-08-475-985-48 Sequence 48, Appl
37 80 8.4 53 3 US-08-256-839-48 Sequence 48, Appl
38 80 8.4 55 1 US-08-475-989-47 Sequence 47, Appl
39 80 8.4 55 2 US-08-475-985-47 Sequence 47, Appl
40 80 8.4 55 3 US-08-256-839-47 Sequence 47, Appl
41 164 4 US-09-370-838-203 Patent No. 5262177
42 80 8.4 696 6 5262177-4 Patent No. 5262177
43 80 8.4 698 2 US-08-175-158A-2 Sequence 2, Appl
44 80 8.4 1074 2 US-08-470-058-2 Sequence 2, Appl
45 80 8.4 1074 3 US-09-037-188-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-472-172-4
; Sequence 4, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y
; APPLICANT: Fahlm, Raafat
; APPLICANT: Sla, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slm & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-472-172-4

		Query Match	58.4%;	Score 559;	DB 2;	Length 455;
		Best Local Similarity	61.6%;	Pred. NO.	I.7e-52;	
		Matches 106;	Conservative 29;	Mismatches 31;	Indels 6;	Gaps
QY	I2	KGSLTLKPAYNELSGFHQLTQWAIHYSKYKTWSRFKELRGKYQDGSYGAFYTKKEY	71			
		: : : : : :	:			
Dd	286	KGNLTLLDPYLELSCGFHQLTDFKAHVHSYKTYTHWSRLTKLHASPEDGK - KNPDKELQY	343			
QY	72	KNSRPAITTTVSLNDALTFRAGLAYDKAASTKTHLSASIPDTDRMWSIGATYKFTPNLS	131			
		: : : : : : : : : :	:			
Dd	344	SNNSRVALCASYNLEKLTRLAGIAYDQAASHRSAAIPDTRTWYSLGATYKFTPNLS	403			
QY	132	VDVGFAPHLRGKKHPFETONIKG----LLLLVDATTTCATANLYGLNLNYRF	179			
		: : : : : : : : :	:			
Dd	404	VDLGVAYLKGKGVHKFEKVNTIGDKRTLTTANTYTTSOAHANLYGLNLYSF	455			

```

RESULT 3
US-08-472-172-6
; Sequence 6, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pei Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; OF HAEMOPHILUS INFLUENZAE TYPE B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release 1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-6

```

DB 341 SNNRSLALGASYNLDEKLTLAGIAYDQASRHRSAAIPTDNRWTYSLGATYKFTPNLS 400
QY 132 VDVGFALRGKKKKHFVETONIK-GLLLVEADY 162
DB 401 VDLGYAYLKGKKVHFKEAQOAGGFITTTANY 432

RESULT 4

US-08-472-172-12
; Sequence 12, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVertry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,172
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/849,411
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973

REFERENCE/DOCKET NUMBER: 1038-471
TELEPHONE: (416) 595-1155
TELEX: 063-24567 SIMBAS

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-472-172-12

Query Match 26.3%; Score 252; DB 2; Length 77;
Best Local Similarity 64.9%; Pred. No. 2.9e-20;
Matches 50; Conservative 9; Mismatches 14; Indels 4; Gaps 1;

QY 107 SASIPDTRMXYSIGATYFTPNLSVDVGFALRGKKKKHFVETONIKG---LLLYEADY 162
DB 1 SAAIPDTRMXYSIGATYFTPNLSVDVGFALRGKKKKHFVETONIKG---LLLYEADY 162

QY 163 TTKATANLYGLNLNYRF 179
DB 61 TSOAHANLYGLNLNYSF 77

RESULT 5

US-08-475-989-8
; Sequence 8, Application US/08475989
; Patent No. 5679352
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: Sia, Charles
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELEPHONE: (416) 595-1155
TELEX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-475-989-8

Query Match 13.5%; Score 129; DB 1; Length 34;
Best Local Similarity 66.7%; Pred. No. 2.1e-07;
Matches 22; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 ELSGFHQLTDQWAIHYSYKYTWRSFKELRGKY 56
DB 2 ELSGFHQLTDQWAIHYSYKYTWRSFKELRGKY 56

RESULT 6

US-08-475-985-8
; Sequence 8, Application US/08475985
; Patent No. 5972349
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: Sia, Charles
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae

;; TITLE OF INVENTION: Conjugate Vaccine
;; NUMBER OF SEQUENCES: 56
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: Suite 701, 330 University Avenue
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R7
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/475,985
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/256,839
;; FILING DATE: 03-FEB-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/CA93/00041
;; FILING DATE: 03-FEB-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9202219.3
;; FILING DATE: 03-FEB-1992
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 34 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-475-985-8
Query Match 13.5%; Score 129; DB 2; Length 34;
Best Local Similarity 66.7%; Pred. No. 2.1e-07;
Matches 22; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 24 ELSGFHQLTDQWAIHYSKYKTEWSRPFKELRGKY 56
|||||:|||||:|
Db 2 ELSGFHQLTDKLVHYSKYKTHWSRLTKLHASF 34
RESULT 7
US-08-472-172-20
;; Sequence 20, Application US/08472172
;; Patent No. 5985288
;; GENERAL INFORMATION:
;; APPLICANT: Munson, Jr., Robert S
;; APPLICANT: Grass, Susan
;; APPLICANT: Chong, Pele Y
;; APPLICANT: Fahim, Raafat
;; APPLICANT: Sla, Charles D Y
;; APPLICANT: McVerry, Patrick
;; APPLICANT: Klein, Michel
;; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
;; NUMBER OF SEQUENCES: 26
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: Suite 701, 330 University Avenue
;; CITY: Toronto

;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R6
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/472,172
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/849,411
;; FILING DATE: 07-JUL-1992
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Stewart, Michael I.
;; REGISTRATION NUMBER: 24973
;; REFERENCE/DOCKET NUMBER: 1038-471
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; TELEX: 063-24567 SIMBAS
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 34 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-472-172-20
Query Match 13.5%; Score 129; DB 2; Length 34;
Best Local Similarity 66.7%; Pred. No. 2.1e-07;
Matches 22; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 24 ELSGFHQLTDQWAIHYSKYKTEWSRPFKELRGKY 56
|||||:|||||:|
Db 2 ELSGFHQLTDKLVHYSKYKTHWSRLTKLHASF 34
RESULT 8
US-08-256-839-8
;; Sequence 8, Application US/08256839
;; Patent No. 6018019
;; GENERAL INFORMATION:
;; APPLICANT: CHONG, Pele
;; APPLICANT: KANDIL, Ali
;; APPLICANT: SIA, Charles
;; APPLICANT: KLEIN, Michel
;; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
;; NUMBER OF SEQUENCES: 56
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: Suite 701, 330 University Avenue
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R7
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/256,839
;; FILING DATE:
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb

RESULT 11
US-08-472-172-22
; Sequence 22, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-172-22
Query Match 11.6%; Score 111; DB 2; Length 32;
Best Local Similarity 71.0%; Pred. No. 1.7e-05;
Matches 22; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 85 LNDALTLAGLAYDKAASKTHLSASIPDTR 115
Db 1 LYEKLTLRAGIAYDQAASHRRSAIPDTR 31
RESULT 12
US-08-256-839-10
; Sequence 10, Application US/08256839
; Patent No. 6018019
; GENERAL INFORMATION:
; APPLICANT: Chong, Pele
; APPLICANT: RANDILL, All
; APPLICANT: Sia, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate vaccine

; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,839
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-256-839-10
Query Match 11.6%; Score 111; DB 3; Length 32;
Best Local Similarity 71.0%; Pred. No. 1.7e-05;
Matches 22; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 85 LNDALTLAGLAYDKAASKTHLSASIPDTR 115
Db 1 LYEKLTLRAGIAYDQAASHRRSAIPDTR 31
RESULT 13
US-08-810-655A-2
; Sequence 2, Application US/08810655A
; Patent No. 5948412
; GENERAL INFORMATION:
; APPLICANT: Murphy, Timothy F.
; TITLE OF INVENTION: Vaccine for Moraxella catarrhalis
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods &
; ADDRESSEE: Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.655A
; FILING DATE: 03/03/1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial No. 5948412 08/245.758
; FILING DATE: 17/05/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Kadle, Ranjana
; REGISTRATION NUMBER: 40,041
; REFERENCE/DOCKET NUMBER: 11520.0091
; TELECOMMUNICATION INFORMATION:


```

; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 2 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; OTHER INFORMATION: leader peptide amino acids 1-25
US-08-810-653A-2

Query Match 9.3%; Score 89; DB 2; Length 459;
Best Local Similarity 23.2%; Pred. No. 0.18;
Matches 42; Conservative 19; Mismatches 84; Indels 36; Gaps 6;

QY 9 YIGKGSITLKLPAWELSGPHQLTDQWAIHYSKYTWSRFKELRGYQDGSYEAPTKK 68
DB 292 YAGKDKFKVTLPSWNLDFQTVNPTTLTAKVRYVPWSDFDIRPTQYTE-----TTK 344
QY 69 EEE-----KNSRFAIGTYSLNDALTLRAGLAYDKAASKTHLSASIPDTRMWY 118
DB 345 LRPQGLPIISYDKDQWSAEVGLKRVSDRLAVSGVGDWGAGN---PASSLGPYKGY 401
QY 119 S--IGATYKFTPNLSVDVGFALHGRKKHFVETQNIKGLLLVADYTTKATANLYGLN 176
DB 402 SLGLGARYNVTPEWSLSLG-----GKYFKFGDAQ-----AQLPTKDKVGNFDSNDG 447
QY 177 Y 177
DB 448 Y 448

RESULT 14
US-08-475-989-14
; Sequence 14, Application US/08475989
; Patent No. 5679352
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,989
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MIS:v9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163

```

```

; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-475-989-14

Query Match 9.1%; Score 87.5; DB 1; Length 35;
Best Local Similarity 54.3%; Pred. No. 0.0069;
Matches 19; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 146 FVETQNIK-GLLLVADYTTKATANLYGLNLYRF 179
DB 1 FKEAQAAGGFITTTANTYTSQAHANLYGLNLYSF 35

RESULT 15
US-08-475-985-14
; Sequence 14, Application US/08475985
; Patent No. 5972349
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,985
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163

```


STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/355,844
FILING DATE: 14-DEC-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..340
OTHER INFORMATION: OmpF porin protein
US-08-355-844-1

Query Match 8.9%; Score 85.5; DB 2; Length 340;
Best Local Similarity 25.2%; Pred. No. 0.28; Mismatches 60; Indels 25; Gaps 6;
Matches 36; Conservative 22; Mismatches 60; Indels 25; Gaps 6;
QY 4 GVLGPYIGKSLTLKLPAYWELSGFHQLTDNA-----IHYSKYKTWNSRFKELRG 54
DB 186 GIVGAY-GAARTNLQEA--QPLNGKKAQWATGLKYDANNIYLAANYGETRNATPITN 242
QY 55 KYODSGCYEAFKKEEYKNSRFAICTTYSLNDAITLRLAGLAYDKAASKTHLSASIPDPTD 114
DB 243 KFTNSGFAKQDVLVLAQYQFOFG-----LRPSIATYKSKAKD--VEGICDND 290
QY 115 RM-WYSIGATYKFTPNLSVDVGV 136
DB 291 LVNYFEVGATYFKNKMSYVOY 313

RESULT 19
PCT-US95-16126-1

Sequence 1, Application PC/TUS9516126
GENERAL INFORMATION:
APPLICANT: Fischbarg, Jorge
APPLICANT: Czegledy, Ferenc
APPLICANT: Iserovich, Pavel
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Ononhue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-OOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/355,844
FILING DATE: 14-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..340
OTHER INFORMATION: OmpF porin protein
PCT-US95-16126-1

Query Match 8.9%; Score 85.5; DB 5; Length 340;
Best Local Similarity 25.2%; Pred. No. 0.28; Mismatches 60; Indels 25; Gaps 6;
Matches 36; Conservative 22; Mismatches 60; Indels 25; Gaps 6;
QY 4 GVLGPYIGKSLTLKLPAYWELSGFHQLTDNA-----IHYSKYKTWNSRFKELRG 54
DB 186 GIVGAY-GAARTNLQEA--QPLNGKKAQWATGLKYDANNIYLAANYGETRNATPITN 242
QY 55 KYODSGCYEAFKKEEYKNSRFAICTTYSLNDAITLRLAGLAYDKAASKTHLSASIPDPTD 114
DB 243 KFTNSGFAKQDVLVLAQYQFOFG-----LRPSIATYKSKAKD--VEGICDND 290
QY 115 RM-WYSIGATYKFTPNLSVDVGV 136
DB 291 LVNYFEVGATYFKNKMSYVOY 313

RESULT 20

US-08-961-083-152
Sequence 152, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Chol et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antlgens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-152.

Query Match 8.7%; Score 83.5; DB 4; Length 261;
Best Local Similarity 23.9%; Pred. No. 0.32;
Matches 38; Conservative 26; Mismatches 68; Indels 27; Gaps 7;
QY 8 PYIGKGSILTKLPAYWELSGFHLTDQWATHYSYK----YTWESRFKELRGKYDGSYEA 64
DB 8 PPSKSGTIGPYNGIRLEGPCK--SEW--YFDKNGVLQEFVGVKTLKTKDSVGRKY 63
QY 65 FTKKEYKDNRSFAIGTYSINDALTLAGLAYDKA----ASKTHLS-----ASI 110
DB 64 GEKREDSERKERYNYFNQNHSLTGWLYDOSNNWYLAKEINGENYLGERRAGW 123
QY 111 PTDTRMWYSIGATYKFTPNLSVDVGFHLRGKKKHVFET 149
DB 124 INDOSTWYLDPTTGI-----MOTGWOL-GNKWYLYRS 156

RESULT 21

US-08-183-214-2
Sequence 2, Application US/08I832I4
Patent No. 5716816
GENERAL INFORMATION:
APPLICANT: Moss, Joel
APPLICANT: Stanley, Sally J.
APPLICANT: Nightingale, Maria S.
APPLICANT: Murtagh, Jr., James J.
APPLICANT: Monaco, Lucia
APPLICANT: Takada, Tatsuyuki
TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
ADP-RIBOSYLARGININE HYDROLASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/183,214
FILING DATE: 14-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,231
FILING DATE: 22-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-27
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-183-214-2.

Query Match 8.7%; Score 83.5; DB 1; Length 362;
Best Local Similarity 26.7%; Pred. No. 0.5;
Matches 36; Conservative 19; Mismatches 55; Indels 25; Gaps 7;
QY 11 GKGSILTKLP---AYWELSGFHLTDQWATHYSYKYTEWSRFKELRGKYDGSYEAFTK 67
DB 197 GKGLMEV-LPEAKAYVTQSGY--FVKENLQWVSFEWEKYLELRG-ILDGKSAPIVPO 252
QY 68 KEYKDNRSFAIGTYS-----LNDALTLAGLAYDKAASKTHLSASIPD 114
DB 253 PFGVKERDQFVIEVSYSGWSSGSHDAPMIAYDAL-LAAGDSNKLALRAFFHGGSDST 311
QY 115 R----MWYSIGATYK 125
DB 312 ATTAGCWGMVHGFK 326

RESULT 22

US-08-475-989-49
Sequence 49, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
CONJUGATE VACCINE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155

; CITY: Boston

```

; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,719
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/090,526
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/629,602
; FILING DATE: 18-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/192002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-715-5

Query Match      8.6%; Score 82; DB 1; Length 188;
Best Local Similarity 23.7%; Pred. No. 0.29;
Matches 45; Conservative 25; Mismatches 52; Indels 68; Gaps 11;
US-08-486-715-5

Qy 32 TDQWAIHYS-YKYTEMSRFEKLRG-----KYQDGSYGAEFTK----- 67
Db 25 TNAFSGYARYAQSKVQDFKNIRGVNVKRYEDDSPVFSISLYLYGDRQASGSVEPEG 84
Qy 68 -----KEEKDNRFRAGTGYTSLNDALTL--RAGLAYDKAASKTHLSASIPDTD----- 114
Db 85 IHVHDFEVKXGS-LMVGPAYRLSDNFSLYALAGVGTVKATFKEH---STQDGSFSNKI 140
Qy 115 ---RMVYSICATYKFTP--NLSVDVGFAGHLRCKKHFVETQNIKGLLVEADYTTKATAN 169
Db 141 SSRKTGFAGAGVQMPLENIVVDVG-----EGSNIS-----STK--IN 178
Qy 170 LYGILNLYRF 179
Db 179 GFNVGVGYRF 188

RESULT 26
US-08-486-719-5
; Sequence 5, Application US/08486719
; Patent No. 5674736
; GENERAL INFORMATION:
; APPLICANT: Miller, Samuel I.
; TITLE OF INVENTION: SALMONELLA VIRULENCE GENES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P. C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/090,526
; FILING DATE: 09-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/629,602
; FILING DATE: 18-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/192004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-749-5

Query Match      8.6%; Score 82;
Best Local Similarity 23.7%; Pred. No. 0;
Matches 45; Conservative 25; Mismatch

Qy 32 TDQWAIHYS-YKTEWRRFKELRG-----KYQDGS
Db 25 TNAFSGVARYAQSQVDFKRIQNVNKKYRYEDDS

Qy 68 ----KEFYKDNSRFAIGTYSLNDAITL--RAGL
Db 85 IHVHDKFEVXKGS-LMVGPAYRLSDNFSLYALAGV

Qy 115 ---RMWYSIGATYKFTP--NLVDVGVGAHLRGKKK
Db 141 SSRKTGFAGWAGVQMNPLENIVVDVG-----

Qy 170 LYGLNLNRYF 179
Db 179 GFNVGVGYRF 188
      : : : : :
      : : : : :

RESULT 29
US-08-183-214-10
; Sequence 10, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING HUMAN
; TITLE OF INVENTION: ADP-RIBOSYLARGININE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Str
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Vers
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/888,231
;; FILING DATE: 22-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weber, Kenneth A.
;; REGISTRATION NUMBER: 31,677
;; REFERENCE/DOCKET NUMBER: 15280-27
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-543-9600
;; TELEFAX: 415-543-5043
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 362 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-183-214-10

Query Match 8.5%; Score 81.5; DB 1; Length 362;
Best Local Similarity 27.2%; Pred. No. 0.83;
Matches 37; Conservative 18; Mismatches 54; Indels 27; Gaps 8;
QY 11 GKGSUTLKLPL---AYWELSGFHLTDQWAIHYSKYKTEWSRPEKELRGKYQDGSYEAFTK 67
DB 197 GKGLMEV-LPEAKKVIYTSQGY--FVKENLQHSYEKEKEKYLELRG-ILDONSAPVFPQ 252
QY 68 KEEYKDNRSRAFTGTYTS-----LNDALTLAGLAYDKAAKSTHLSASIPDTD 114
DB 253 PFGVKRQDFYIDVSYSQWSSGSHDPMIAYDAL-LAAGDSWKELAHRAFPFHGSDSDST 311
QY 115 R----MWYSIGATYKF 126
DB 312 AAIAGCWM--GVMYGF 325

RESULT 30
US-08-286-767-3
; Sequence 3, Application US/08286767
; Patent No. 5733760
; GENERAL INFORMATION:
; APPLICANT: Lu, Yichen
; APPLICANT: Miller, Samuel I.
; APPLICANT: Killeen, Kevin
; TITLE OF INVENTION: SALMONELLA VECTORS ENCODING TRUNCATED
; TITLE OF INVENTION: P9c FUSION PROTEIN, METHOD OF MAKING, AND USES THEREO
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RONALD I. EISENSTEIN; DIKE, BRONSTEIN,
; ADDRESSEE: ROBERTS & CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,767
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 44420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6400
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 188 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-286-767-3
Query Match 8.5%; Score 81; DB 1; Length 188;
Best Local Similarity 23.7%; Pred. No. 0.37;
Matches 45; Conservative 25; Mismatches 52; Indels 68; Gaps 11;
QY 32 TDQWAIHYS-YKYTEWSRPEKELRG-----KYDGSYEAFTKKEEYKDNRSRA----- 78
DB 25 TNAFSVGVARYAQSKVQDFKNIRGVNVKRYEDDSPV-SFISLSLYLDGDRQAGSVEPE 83
QY 79 -----IGTYSLNDALTL--RAGLAYDKAAKSTHLSASIPDTD----- 114
DB 84 GIHYHDKPEVKYGLMVGCPAYRLSDNFSLYALAGVTYKATEKEH---STQGDSEFSNKI 140
QY 115 ---RMWYSIGATYKFTP--NLSVDVGFALHKGKKHFEVTONIKGLLLVEADYTTKATAN 169
DB 141 SSRKTGFAMGAGVQMNPLENIVDVGY-----EGSNIS-----STK--IN 178
QY 170 LYGLNLNRYF 179
DB 179 GFNVGVGYRF 188
RESULT 31
US-08-475-989-13
; Sequence 13, Application US/08475989
; Patent No. 5679352
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,989
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA93/00041
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9202219.3
; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-475-989-13

Query Match 8.4%; Score 80; DB 1; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 160 ADYTKATANLYGLNLYRF 179
|:|:|:| | | | | | | | | | |
DB 19 ANYTQAHANLYGLNLYSF 38

RESULT 32
US-08-475-985-13
; Sequence 13, Application US/08475985
; Patent No. 5972349
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,985
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA93/00041
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9202219.3
; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-475-985-13

Query Match 8.4%; Score 80; DB 2; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 160 ADYTKATANLYGLNLYRF 179
|:|:|:| | | | | | | | | | |
DB 19 ANYTQAHANLYGLNLYSF 38

RESULT 33
US-08-472-172-25
; Sequence 25, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: SIA, Charles D Y
; APPLICANT: McVerry, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein p1 and Peptides
; TITLE OF INVENTION: of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELETYPE: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-172-25

Query Match 8.4%; Score 80; DB 2; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 160 ADYTKATANLYGLNLYRF 179
|:|:|:| | | | | | | | | | |
DB 19 ANYTQAHANLYGLNLYSF 38

RESULT 34
US-08-256-839-13
; Sequence 13, Application US/08256839
; Patent No. 6018019

GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256.839
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-839-13

Query Match 8.4%; Score 80; DB 3; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 160 ADYTKATANLYGLNLNRYF 179
|:::| | | | | | | | | |
Db 19 ANYTSQAHANLYGLNLNYSF 38

RESULT 35
US-08-475-989-48
Sequence 48, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475.989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256.839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-989-48

Query Match 8.4%; Score 80; DB 1; Length 53;
Best Local Similarity 75.0%; Pred. No. 0.081;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 160 ADYTKATANLYGLNLNRYF 179
|:::| | | | | | | | | |
Db 34 ANYTSQAHANLYGLNLNYSF 53

RESULT 36
US-08-475-985-48
Sequence 48, Application US/08475985
Patent No. 5972349
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475.985
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256.839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041

;; FILING DATE: 03-FEB-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9202219.3
;; FILING DATE: 03-FEB-1992
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 48:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 53 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-475-985-48

Query Match 8.4%; Score 80; DB 2; Length 53;
Best Local Similarity 75.0%; Pred. No. 0.081;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 160 ADYTKRATANLYGLNLYRF I79
|:::| | | | | | | | | |
Db 34 ANYTQAHANLYGLNLYSF 53

RESULT 37
US-08-256-839-48
;; Sequence 48, Application US/08256839
;; Patent No. 6018019
;; GENERAL INFORMATION:
;; APPLICANT: CHONG, Pele
;; APPLICANT: KANDIL, Ali
;; APPLICANT: SIA, Charles
;; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
;; NUMBER OF SEQUENCES: 56
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: Suite 701, 330 University Avenue
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R7
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/256,839
;; FILING DATE:
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 48:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 53 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-256-839-48

Query Match 8.4%; Score 80; DB 3; Length 53;
Best Local Similarity 75.0%; Pred. No. 0.081;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 160 ADYTKRATANLYGLNLYRF I79
|:::| | | | | | | | | |
Db 34 ANYTQAHANLYGLNLYSF 53
RESULT 38
US-08-475-989-47
;; Sequence 47, Application US/08475989
;; Patent No. 5679352
;; GENERAL INFORMATION:
;; APPLICANT: CHONG, Pele
;; APPLICANT: KANDIL, Ali
;; APPLICANT: SIA, Charles
;; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
;; NUMBER OF SEQUENCES: 56
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: Suite 701, 330 University Avenue
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R7
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/475,989
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/256,839
;; FILING DATE: 03-FEB-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/CA93/00041
;; FILING DATE: 03-FEB-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9202219.3
;; FILING DATE: 03-FEB-1992
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 47:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 55 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-475-989-47

Query Match 8.4%; Score 80; DB 1; Length 55;
Best Local Similarity 75.0%; Pred. No. 0.085;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 160 ADYTKRATANLYGLNLYRF 179
|:::| | | | | | | | | |
Db 36 ANYTQAHANLYGLNLYSF 55

RESULT 39
US-08-475-985-47
; Sequence 47, Application US/08475985
; Patent No. 5972349
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475.985
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256.839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA93/00041
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9202219.3
; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 55 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-475-985-47

Query Match 8.4%; Score 80; DB 2; Length 55;
Best Local Similarity 75.0%; Pred. No. 0.085;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 160 ADYTTKATNANLYGLNLYRF 179
|:::| | | | | | | | | |
Db 36 ANYTSQAHANLYGLNLYSF 55

RESULT 40
US-08-256-839-47
; Sequence 47, Application US/08256839
; Patent No. 6018019
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel

; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256.839
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 55 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-256-839-47

Query Match 8.4%; Score 80; DB 3; Length 55;
Best Local Similarity 75.0%; Pred. No. 0.085;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 160 ADYTTKATNANLYGLNLYRF 179
|:::| | | | | | | | | |
Db 36 ANYTSQAHANLYGLNLYSF 55

Search completed: May 12, 2003, 09:41:59
Job time : 17 secs